

DOCUMENT RESUME

ED 120 586

CE 006 997

TITLE Field-Based Testing of the Ohio Career Development Program.

INSTITUTION Policy Studies in Education, New York, N.Y.

REPORT NO VT-102-762

PUB DATE Jul 75

NOTE 122p.; Not available in hard copy due to marginal reproducibility of the original

EDRS PRICE MF-\$0.83 HC-\$6.01 Plus Postage

DESCRIPTORS Affective Tests; *Career Education; Cognitive Tests; Comparative Analysis; *Curriculum Evaluation; Elementary Secondary Education; Evaluation Methods; Participant Characteristics; *Program Effectiveness; Program Evaluation; *State Programs; State Schools; *Student Testing; Tables (Data); Testing; Test Results; Vocational Development

IDENTIFIERS *Ohio; Program Nonparticipants

ABSTRACT

The report presents the fourth evaluation of the Ohio Career Development Program for Grades K-10. Although career education programs appeared to be successful on observation, the administration of curriculum-based tests in 1973 showed no major difference between results of the program and nonprogram students. Policy studies in education developed and administered field-based tests in order to test the hypothesis that program students were learning more than nonprogram students, but that program students could not fully demonstrate their knowledge on curriculum-based tests. The outcome of this project resulted in a set of test results for grades 3, 6, 8, and 10 in seven cities, and a set of test instruments for these grades with potential for statewide use. Program students did score significantly higher than nonprogram students on both the cognitive and affective tests. The field-based method of creating tests successfully identified what program students were learning as well as what they were not learning. The tests also told in what ways program students were superior to, equal to, and inferior to nonprogram students. The report details the background of the study, methodology, outcomes of the test development procedure, and findings. Thirty-five of the tabulated test results are appended. (Author/RG)

* Documents acquired by ERIC include many informal unpublished *
* materials not available from other sources. ERIC makes every effort *
* to obtain the best copy available. Nevertheless, items of marginal *
* reproducibility are often encountered and this affects the quality *
* of the microfiche and hardcopy reproductions ERIC makes available *
* via the ERIC Document Reproduction Service (EDRS). EDRS is not *
* responsible for the quality of the original document. Reproductions *
* supplied by EDRS are the best that can be made from the original. *

ED120586

FIELD-BASED TESTING OF THE OHIO CAREER DEVELOPMENT PROGRAM

BEST COPY AVAILABLE

POLICY STUDIES IN EDUCATION

52 VANDERBILT AVENUE
NEW YORK, NEW YORK

A DEPARTMENT OF THE EDUCATIONAL RESEARCH COUNCIL OF AMERICA

JULY, 1975

VT102762

2

U.S. DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

CE006987

Table of Contents

	<u>Page</u>
ACKNOWLEDGMENTS	1
BACKGROUND	1
METHODOLOGY	15
OUTCOMES OF THE TEST DEVELOPMENT PROCEDURE	36
FINDINGS	58
APPENDICES	

ACKNOWLEDGEMENTS

PSE began working with the staff of the Ohio Career Development Program in 1972. Since that time, Dr. Byrl Shoemaker, Director, Division of Vocational Education, has taken an extremely active interest in our activities and has firmly supported our evaluation efforts. This year was no exception. We extend our appreciation to Dr. Shoemaker for his continuing support.

Jack Ford, Assistant Director, Division of Vocational Education, Career Development Service, assisted us this year in innumerable ways. Always responsive to our requests for information, help, or advice, he helped make this year's evaluation run smoothly in every respect.

We also want to thank Susan Sears, Supervisor, Career Development Service, who contributed many thoughtful suggestions during the year.

The Program Directors in the seven participating school districts deserve special thanks for their assistance in making arrangements for item writing, pilot testing, and final testing. The Program Directors offered the help of their building Coordinators at many points in the study in assisting with field logistics, explaining the test procedures to classroom teachers, administering the tests on some occasions, supplying background information about their schools, and so on.

The Program Directors arranged for access to classrooms for observations, interviews with teachers, pilot test administration and field test administration, duplication services, and more. They completed numerous forms and questionnaires and reviewed test items on request. They carried out each assignment quickly and graciously.

We would like to single out for special mention Nicholas Topouguis, Program Director in Akron, and Dorothy Riley and Margaret Busler, Program Directors in Mansfield, both in the midst of the new responsibilities this year, for special thanks.

Donald Burch, Program Director in East Muskingum and Warren Doerr, Program Director in South-Western City deserve special mention as well.

We are also grateful to Marlyn Bray, Program Director in Cincinnati; George Degenhart, Program Director in Springfield; and Oliver Jones, Program Director in Cleveland for their assistance.

Finally, our thanks go to all the classroom teachers and their students who took part in our activities. Their cooperation and participation were at the heart of the year's work. They made it all possible. To them we are indeed grateful.

BACKGROUND: 1972-1974

The Ohio Career Development Program has been under evaluation for several years. This report can be best understood in its historical context.

The Ohio Career Development Program

The present program emerged over a number of years, changing steadily from a set of separate projects at the beginning into an integrated program today. The transition from separate projects into an integrated program is reviewed briefly below.

The Start of Career Education in Ohio. Career education began in Ohio in the late 1960s when the State Department of Education through the Division of Vocational Education funded local school districts to operate individual career education projects designed to serve selected grade ranges. One school district, for example, decided to start by operating a career education project for students in grades K-6--a project called the World of Work. A second district began with a project for junior high school students in grades 7-8--a project called Career Orientation. Some districts began high school projects.

Designing the Ohio Career Development Program for Grades K-10.

In 1971, the Division decided to relate and extend all local projects so as to provide a continuous career education experience for students in grades K-10 in participating school districts. To accomplish that, the Division developed a set of State Curriculum Guides and asked all funded districts to use them.

The comprehensive Ohio Career Development Program which resulted

was first implemented in 1972-73. It was comprised of the following segments, each of which became an integral part of the program: Career Motivation (originally called the World of Work) for grades K-6; Career Orientation for grades 7-8; and Career Exploration for grades 9-10.

Number of School Districts Participating. In 1972-73, there were twenty districts participating in the program; by 1973-74, twenty-four districts; and by 1974-75, thirty districts.

Evaluating the Ohio Career Development Program

The program has been under intensive evaluation for three years. During that time, four major evaluations have been conducted by staff from Policy Studies in Education (PSE):

1. The 1972 Observation/Evaluation
2. The 1973 Development and Administration of Curriculum-Based Tests
3. The 1973-74 Evaluation/Observation
4. The 1974-75 Development and Administration of Field-Based Tests.

The following pages summarize the first three evaluations and explain how they led to the fourth. The fourth evaluation--conducted in 1974-75--is described in complete detail later; it constitutes the main body of this report.

The 1972 Observation/Evaluation

In 1972-73, with twenty districts participating, staff from PSE were asked by the Division of Vocational Education to conduct site visits and interviews in all twenty districts in order to answer the following four questions:

1. Is career education taking place in the classrooms?
2. Is an average allotment of \$25.00 per student enough money to implement career education in the twenty participating districts?
3. Is an average allotment of \$25.00 per student enough money to spread career education throughout the State of Ohio?
4. Should the career development program continue to be administered by the Division of Vocational Education?

Career Education Appeared To Be Successful. The report, An Evaluation of the Ohio Career Development Program, was submitted in February, 1973. Summarized answers to the four questions, which were elaborated in the report, are as follows:

Question 1: Is career education taking place in the classrooms?

"In summary, career education in Ohio is beginning to take place in classrooms, especially in grades K-6, but in all grades it appears chiefly as additional activities rather than as fully integrated parts of existing curricula."

Question 2: Is an average allotment of \$25.00 per student enough money to implement career education in the twenty participating districts?

"The average allocation of \$25.00 per child is enough money to implement the Career Development Program in the twenty participating school districts--implement it to the degree indicated in the answer to Question 1 above."

Question 3: Is an average allotment of \$25.00 per student enough money to spread career education throughout the State of Ohio?

"Generally, it can be said that since \$25.00 per child is enough to implement a Career Development Program in twenty districts, the same

allotment could suffice for other districts through the State which have characteristics in common with the present twenty."

Question 4: Should the Career Development Program continue to be administered by the Division of Vocational Education?

"Particularly during the early years of the Career Development Program and perhaps for years to come, career development can find no better administrative home within the Ohio Department of Education than it has found in the Division of Vocational Education."

The report also contained a series of recommendations for strengthening and improving the program. Some recommendations specifically related to the four questions; others reflected additional observations as a natural by-product of the visits and interviews.

The 1973 Development and Administration of Curriculum-Based Tests

In February, 1973, the Division asked PSE project staff to conduct an evaluation of the Career Development Program through testing students in four of the twenty school districts.

The four sites chosen by the Division for that evaluation were the four required to have an external evaluation as a condition of Federal funding: Ohio's one VEA Part C site and its three VEA Part D sites. The four sites constituted a reasonably representative cross-section of the kinds of schools to be found in most of the twenty school districts then taking part in the program. Thus the Division anticipated that by using the Part C and Part D sites as the locations for evaluating student learning, whatever it learned could be applied to improving the program in most of the other school districts.

An evaluation plan was proposed with two major objectives:

1. To develop and pilot test a sequential battery of cognitive and affective instruments for grades 3, 6, 8, and 10 based on the Career Development Program State Curriculum Guides issued for grades K-6, grades 7-8, and grades 9-10. The tests would be designed to match the terminal points of the major program segments (grades K-3, 4-6, 7-8, and 9-10).
2. To administer the final instruments to a representative sample of program and non-program students in the four Part C and Part D sites participating in the Career Development Program in order to evaluate its impact on student learning.

The instruments themselves would become measuring devices for comparing intended student learning to actual student learning, not only in the four initial sites but in additional sites in the future. The information derived from the first administration of the instruments would be reported to the Division so that the state leaders could modify program practices in participating school districts if necessary in carrying out their commitment to a high quality Career Development Program.

The tests were developed in the spring of 1973 and were administered in May, 1973 to students in grades 3, 6, 8, and 10 in the one VEA Part C site and the three VEA Part D sites. The report of the findings, Testing the Ohio Career Development Program, was submitted to the Division in August, 1973.

No Major Differences Found Between Program and Non-program Students.

Both groups scored well on the tests. There were only small differences in favor of program students on both the cognitive and affective tests in all grades--3, 6, 8, and 10. The small differences were less than what the PSE staff expected they would be; they were less than what the

Division hoped they would be. We were not surprised that non-program students could answer some questions, but we were surprised that program students could not answer even more questions, thus which would have demonstrated extra learning attributable to their being in the program.

The PSE staff examined the data thoroughly, looking for patterns that could explain why program students did not score higher than non-program students, or for patterns that might suggest how the program could be improved. While no clear-cut explanations or patterns emerged, a number of possible interpretations of the findings were offered in the final report. Below is a summary of the interpretations written at that time:

Interpretation 1: No Career Education Is Taking Place. It is possible that career education has not yet been effectively implemented in program classrooms. If so, it would mean that no matter what care and expense were put into developing tests, they would show no differences between program and non-program students.

We are inclined to reject this conclusion based on the 1972-73 evaluation of the program in local school districts, where local project leadership activities were highly visible and where there was enough evidence of career education in the classrooms to make PSE observers believe that students were learning.

Interpretation 2: Other School Curricula Teach Career Concepts. The content of the Career Development Program has much in common with other school curricula. That is, the relationship between self and environment, ideas about the world of work, concepts of economics, skills in decision-making, and other areas of concern in career

development also appear in one form or another in many other subject fields. Thus the students in non-program classes have elaborate opportunities to learn that content in the study of various subjects at various grade levels.

While the overlap of career development content with the content of other curricula does not fully explain why program students did not demonstrate extra learning on the tests, it seems reasonable to believe that it helps explain why non-program students did so well.

Interpretation 3: Non-School Sources Teach Career Concepts. There are many ways for students to learn career concepts outside of school: at home, at play, at camp, at church, while reading, while watching television, with friends, with adults, and so on. This is especially true for the important career concepts measured by the tests. Society does not leave those important concepts to chance; it transmits them to young people in many ways. Since students outside the Career Development Program are exposed to these strong non-school influences, they learn enough to score about as well on the tests as program students.

While this interpretation does not account for the absence of any extra career learning by program students, it may explain why non-program students did well on the tests.

Interpretation 4: Program Students Cannot Fully Demonstrate Their Learning on the Tests. The tests reflect the program as specified in the State Curriculum Guides for grades K-6, 7-8, and 9-10. But the tests may be a better fit to the State Guides than to what is actually happening in local program classrooms. Perhaps the program students cannot fully demonstrate on these tests what they have been learning--and thus

they cannot demonstrate their superiority to non-program students. That is, perhaps the tests do not fully reflect all that is being taught and learned in the program classrooms.

Given the fact that local programs often vary significantly from state guidelines, this interpretation seems reasonable, at least as a partial explanation of what is happening.

The 1973-74 Observation/Evaluation

Uncertain which of these four hypotheses was true, the Division asked PSE to observe the ongoing program once again in a small sample of sites. In the spring of 1974, PSE staff visited four program sites selected by the Division. At each site, the staff examined the project for its:

1. Quality of project staff leadership.
2. Teacher use of Career Development Program ideas and materials.
3. Student interest and participation in Career Development Program activities.

PSE staff observed classes, accompanied the teachers and students on field trips, participated in other career-related activities, and conducted interviews with project staff, teachers, and students.

Again Career Education Appeared To Be Successful. While there were some variations among the four sites, the PSE staff found that:

1. Local leadership was generally effective.
2. There was evidence of teacher use of Program ideas and materials in their classrooms. The evidence was greatest at the elementary level, not as great at the junior high level, and weakest at the high school level.

3. Student interest and participation paralleled teacher interest and participation and accordingly was most evident at the elementary level, less evident at the junior high level, and least evident at the senior high level.

On the basis of these findings, PSE rejected Interpretation 1: It concluded that career education was taking place in the program classrooms. We had seen it in the fall of 1972; we had seen it again in the spring of 1974.

The Evidence of the Observations Versus the Evidence of the Instruments

Were we to believe our eyes or our instruments? Based on the new observational evidence that career education had indeed reached the classrooms, PSE believed that Interpretation 4 given earlier was probably correct. That is, students in the program were indeed learning more than other students but it was something that they could not fully demonstrate on the original tests written to match the State Curriculum Guides. (Interpretations 2 and 3--school and non-school sources were teaching much career content to students outside the program--continued to seem quite reasonable as well.)

Field-Based Tests to Supplement Curriculum-Based Tests. PSE believed that the way to verify Interpretation 4 would be to develop field-based tests to match the students' actual learning rather than curriculum-based tests to match the students' intended learning as called for in the State Guides. Presumably, this could be done by going into the field and developing tests on site--close to the classroom where the actual learning was taking place--rather than at an off-site location where the test developers would have nothing but descriptions of intended learning to guide them.

PSE believed that tests developed in this fashion--in and around pro-

gram classrooms in the field--would provide the Division with evidence of career learning by discriminating between program and non-program students and showing the superiority of those in the program. Accordingly, PSE suggested that the Division call for the use of a variety of on-site test development methods in a representative cross-section of school districts as a means of creating cognitive and affective test items which would successfully measure the full extent of student learning in program classrooms. Those test items would then be compiled into paper and pencil tests which presumably would discriminate between program and non-program students in grades 3, 6, 8, and 10.

Retaining the Curriculum-Based Tests. As to the original curriculum-based tests built to match the State Guides, PSE believed that they should be revised for future use by the Division. Properly designed, such tests could constitute a description--more exactly, a sampling, of a description--of intended student learning. They could be used for such important purposes as these:

The test instruments themselves could serve:

1. As an "ideal" standard for what students should learn.
2. As a target toward which teachers could direct their efforts.
3. As a criterion against which to compare the eventual content of the field-based tests.

When administered in the future, the tests could produce data to serve:

1. As a measure of the discrepancy between intended learning and actual learning.
2. As a measure of progress from the baseline established for program and non-program students when the tests were first administered in May, 1973. (Only the original, unrevised items remaining in the test could be used to measure progress, of course.)

The 1974-75 Original Plan for Development and Administration of Field-Based Tests

In response, the Division issued a call for competitive proposals built around the concept of field-based tests. PSE submitted a competitive proposal to the Division. It described a plan for evaluating career education in seven representative school districts and simultaneously producing a test instrument that might be used statewide subsequently to evaluate career education in all of the thirty districts participating in the Career Development Program.

The original plan is shown schematically on the following page. It envisioned three major tasks to be carried out in selected sites from among the seven.

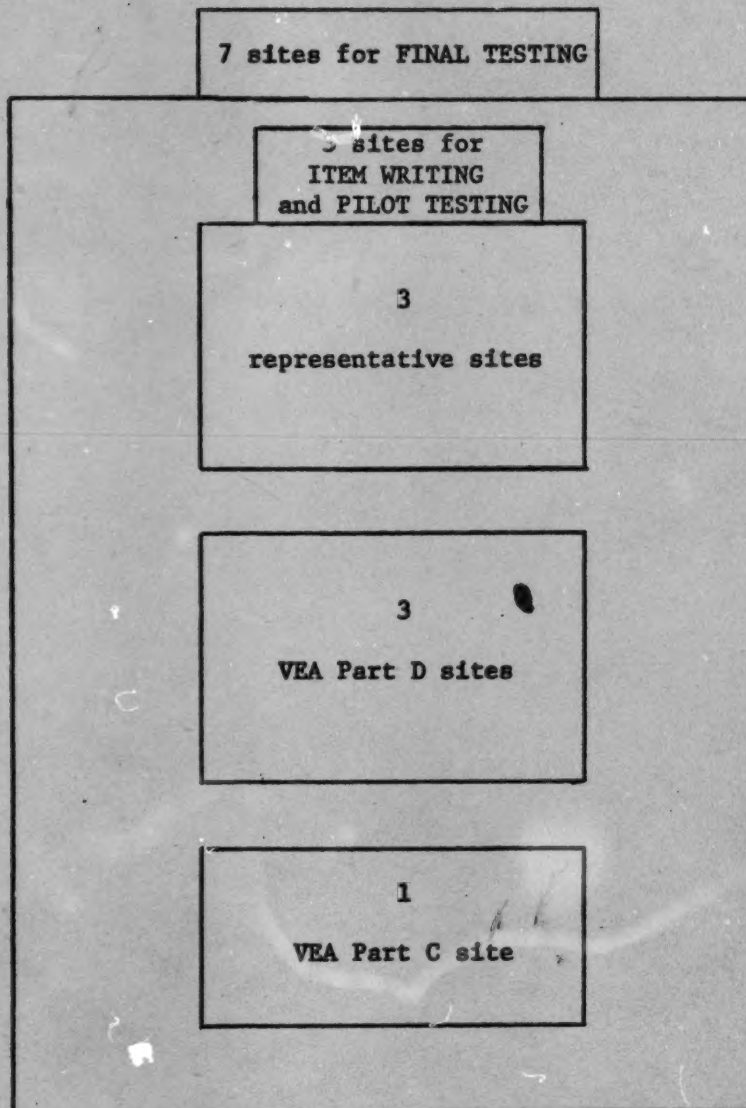
- ITEM WRITING would take place in three sites--one urban, one suburban, one rural--selected by the Division as representative of all thirty sites participating in the program.
- PILOT TESTING of the items would take place in the three item-writing sites.
- FINAL TESTING would take place in the three item-writing sites and in Ohio's three VEA Part D sites and its one VEA Part C site.

The proposal also described the methods to be used in writing the items, pilot testing them, selecting and assembling items which performed satisfactorily in the pilot tests, and administering the final tests.

This plan had several advantages:

- It provided for developing field-based tests in a microcosm of the state (three representative sites) in the expectation that tests proving satisfactory in the three sites would prove satisfactory statewide.
- It provided for administering the final tests in the three VEA Part D sites and the one Part C site. This

1974-75 Original Plan
for Test Development and Administration



would serve not only to try out the battery of instruments in sites different from those in which they were developed, but would also meet the Federal requirement for an external evaluation in the Part D and Part C sites.

- It provided for finding whether the instruments worked better in the sites in which they were developed because the item writers knew more about the program they were testing.

The proposal called for on-site item writing to begin in October, 1974, for pilot testing to take place in the months following, and for final testing to take place in May, 1975.

The Division awarded a contract to PSE in July, 1974, to carry out the proposal. PSE immediately began preparing materials to use in writing items during the field visits beginning in October.

The New USOE Evaluation Guidelines Delay the Plan. In August, 1974, with PSE immersed in its plans for the October visits, the U.S. Office of Education issued its new Draft Guidelines for the Evaluation of Career Education Programs. The Guidelines mandated specific evaluation procedures for all VEA Part D program sites. The procedures were to be implemented during the current 1974-75 school year.

The Guidelines procedures were quite different from the ones PSE proposed to use in Ohio. They called for the administration of designated nationally-validated career education instruments. In contrast, PSE wished to develop new field-based instruments that would reflect Ohio's career education program and which would show differences in learning between program and non-program students.

PSE believed that it was on the brink of a new kind of test development procedure that could produce career education tests capable of uncovering the full range of learning actually taking place; it believed

its test development procedures might provide sorely needed evidence of program impact to inform state and Federal officials interested in career education. Moreover, PSE did not believe that the use of existing instruments could demonstrate that students in career education programs were learning more than other students--either in Ohio or in other sites across the country--since most existing instruments had not been developed specifically with current career education programs in mind and thus might have been insensitive to program effects.

In an attempt to 1) proceed with its experiment in a new kind of test development, and 2) provide an alternative evaluation methodology to the Guidelines should they fail to produce evidence of career education learning, PSE obtained the approval and support of the Ohio Division to meet with USOE officials to request an exception to following the procedures mandated in the Guidelines. In time, USOE officials granted the exception and PSE once again embarked on its plan to develop new and, hopefully, discriminating field-based career education tests. Confering with Division and USOE officials and waiting for the exception to the Guidelines took from September, 1974, to December, 1974.

Revising the Plan. In December, 1974, when PSE received permission to go ahead with its plan in Ohio, it submitted a revised proposal to the Division which modified and compressed some of the elements in its earlier proposal in order to complete in five months the same amount of work originally scheduled for nine months. The Division approved the modified proposal in December, 1974, and PSE set to work in January, 1975.

The remainder of this document describes the methods used and the findings obtained from that point forward.

METHODOLOGY: 1974-75

The revised proposal which PSE submitted to the Division in December, 1974, and which the Division approved in the same month is described below, as modified through subsequent experience in the field. PSE had never tried the approach before and anticipated that a certain amount of experimentation and subsequent modification would be necessary.

Overview of the 1974-75 Revised Plan

The revised plan involved changes in the field locations of the activities, modifications in the procedures, and shifts in the timetable. The revised plan for those arrangements is explained briefly below.

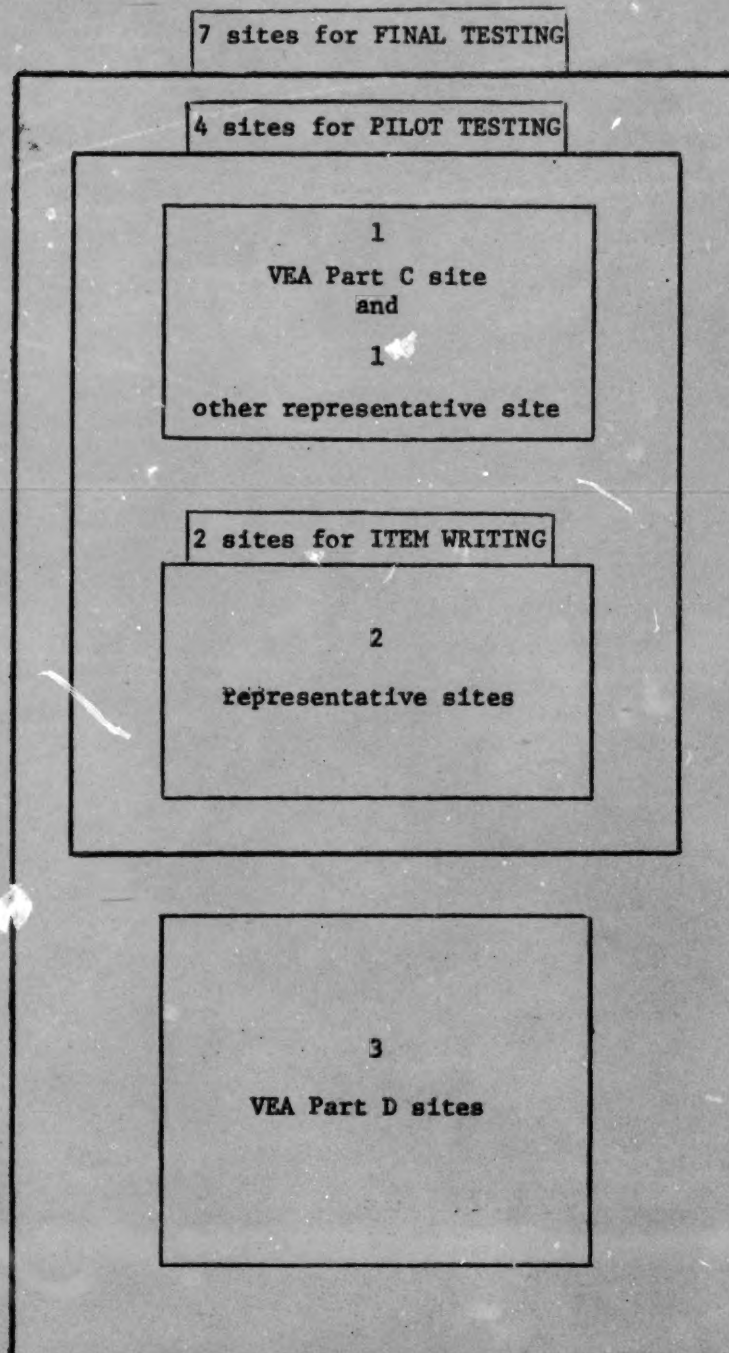
Using Selected Sites for Selected Activities. Compressing the evaluation from nine months to five months required a change in the original plan for writing items and pilot testing items in selected sites from among the seven participating. The new plan is displayed schematically on the following page. It called for:

- ITEM WRITING in 2 sites
- PILOT TESTING in 4 sites, including the two
ITEM WRITING sites
- FINAL TESTING in 7 sites, including the four
PILOT TESTING sites

(The seven sites were Akron, Cincinnati, Cleveland, East Muskingum, Mansfield, South-western City, and Springfield.)

While the two sites selected for item writing were not as fully representative of the state program as the three sites originally selected for item writing, they nevertheless had demographic variations similar

1974-75 Revised Plan
for Test Development and Administration



to those in the state as a whole. Moreover, both sites selected for item writing had high quality career development programs and therefore were rich sources of ideas for test items.

Furthermore, in order to make the final tests as representative as possible, PSE decided to conduct pilot testing in four sites rather than three, as originally planned. That is, by offsetting the loss of one item writing site by adopting an additional pilot testing site, PSE increased the likelihood that the final tests would be sensitive to the full range of career education activities taking place statewide.

The best evidence of whether the final tests were sensitive to program activities statewide would come from administering the final tests not only in the four sites in which items had been pilot tested but also in Ohio's three VEA Part D sites, none of which were involved in the creation of the tests.

Procedures. PSE employed a variety of methods to write items in two sites. Each item was then put into a pilot trial in the two item-writing sites and in two additional sites. It was administered in the program classrooms for which it was written, in other program classrooms, and in non-program classrooms. Those items which discriminated most strongly in favor of program students were compiled into a 60-item final test for each of the grades 3, 6, 8, and 10. Those tests were then administered to program and non-program students in grades 3, 6, 8, and 10 in seven cities: the four pilot testing sites and three additional sites.

The procedures involved these steps:

- | | |
|------------------|--|
| ITEM
WRITING | 1. Preparing the two item-writing sites for the impending on-site activities. |
| | 2. Conducting the on-site item writing. |
| | 3. Polishing and refining the items written on site. |
| | 4. Selecting from the earlier curriculum-based tests any supplementary items which matched local activities. |
| PILOT
TESTING | 5. Administering the items in the four pilot testing sites. |
| | 6. Scoring and analyzing the items to determine which discriminated in favor of program students. |
| | 7. Compiling the discriminating items into a pool for possible inclusion in the final tests. |
| | 8. Compiling the non-discriminating items (plus any which discriminated against the program students) into a pool to send to the Division at the close of the study. |
| | 9. Selecting the most discriminating items for the final tests. |
| FINAL
TESTING | 10. Administering the final tests. |
| | 11. Scoring and analyzing the final tests. |
| | 12. Reporting the results. |

Timetable. The following timetable lists key events in the process as well as indicating when and where they took place.

TIMETABLE

<u>DATE</u>	<u>PLACE</u>	<u>EVENT</u>
1974		
July- Aug.	PSE/New York City	Prepare for test development process.
Aug.	USOE/Washington, D.C.	Publication of <u>USOE Guidelines</u> for the evaluation of career education programs.
Sept.- Dec.	PSE/New York City and USOE/Washington, D.C.	Request exception from <u>Guidelines</u> ; meet with USOE staff; wait for USOE reply.
1975		
Jan. 6-9	Site 1	Item writing I.
Jan. 10	Site 1	Pilot test I.
Feb. 3-7	Site 2, Site 3, Site 4	Pilot test I.
Feb. 3-7	Site 1	Item writing II.
Feb. 10-21	PSE/New York City	Data analysis of pilot test I.
Feb. 17-21	Site 1, Site 2, Site 3, Site 4	Pilot test II.
Feb. 24- March 7	PSE/New York City	Data analysis of pilot test II.
April 14-18	Site 2	Item writing III.
April 28- May 2	Site 1, Site 2, Site 3, Site 4	Pilot test III.
May 5-16	PSE/New York City	Data analysis of pilot test III.
May 19-22	PSE/New York City	Selection of items for inclusion in final tests.
May 27-30	Site 1, Site 2, Site 3, Site 4, Site 5, Site 6, Site 7	Administration of final tests.
June 1-19	PSE/New York City	Data analysis of final tests.
June 20- July 15	PSE/New York City	Preparation of final reports.

Writing the Items

A total of three intensive site visits were conducted to write items: two to Site 1 and one to Site 2.

Preparing for the Visits. Before making the first visit to Site 1, PSE sent the Program Director and his staff the following materials:

1. A site visit schedule
2. A teacher activity guide

Both appear in Appendix A. The schedule and guide were meant to provide the local staff with an orientation to the test development process and to give them enough information to make the necessary arrangements for our visit. For example, PSE specified how many classes at the elementary, junior high, and senior high level the staff wanted to visit, how many students the staff wanted to interview, what materials the staff would need, and so on. The schedule and guide reflected PSE's desire to structure the time on site so as to spend as much time as possible either observing or interviewing students. Further, by asking the project staff to notify teachers of the impending visit, PSE could insure observations of lessons or activities with career content. (PSE assumed that a certain amount of classroom spontaneity would be lost by notifying teachers beforehand, and that the lessons would probably have a somewhat forced, "demonstration-lesson" quality. Nevertheless, PSE thought prior notification to teachers was necessary in order to maximize the value of PSE staff time on-site. In fact, most lessons or activities PSE staff observed turned out to be informal and natural.)

In the subsequent visit to Site 1 and the visit to Site 2, PSE was able to make all arrangements by telephone and letter and did not prepare a schedule or guide.

Item Writing I. A team of two PSE staff members spent January 6 to January 9 in selected Site 1 program schools. The daily log listing the names of the schools, the numbers and grade levels of classrooms they visited, the program staff and teachers they interviewed, and the activities they conducted appears in Appendix B.

Item Writing II. A team of three PSE staff members spent February 3-7 in several Site 1 program schools. The log in Appendix B lists the names of the schools, the number and grade level of classrooms they visited, the program staff and the teachers they interviewed, and the activities they conducted.

During the site visit, PSE staff--sometimes working in pairs, sometimes singly--continued to generate items. As can be seen in the log, some students were interviewed while performing duties in the vocational wing of one high school, but in general staff relied on classroom observation and small group discussions as their main techniques for getting item content.

Item Writing III. A team of two PSE staff members spent April 14-18 working in Site 2. Again, the log in Appendix B describes their day-to-day activities. Both staff members had participated in the Site 1 field visits and had by now refined their item-generating techniques and item-writing skills.

Selecting Supplementary Items From the 1973 Ohio Curriculum-Based Tests. To supplement the items being written in the field in the winter and spring of 1975, PSE turned to the Ohio Career Development Tests that it had produced in 1973 to match the Ohio State Curriculum Guides. The items

in that test were known to be 1) technically sound and 2) designed to reflect the Career Development Program. Despite the fact that these test items had not successfully discriminated between program and non-program students when first used, PSE thought it possible that the cumulative effect of the program on student learning in the succeeding two years might enable program students to demonstrate their superiority to non-program students by June, 1975 -- especially if there were evidence that the item content was being taught during the 1974-75 school year. Accordingly, the PSE staff who were conducting the 1975 site visits reviewed the 1973 curriculum-based tests for items that reflected the student learning they were currently observing. Items addressed to current learning were incorporated into the pilot test instruments.

Categorizing the Items by Developmental Area.* When writing items in the field, the PSE staff did not consider to which of the seven developmental areas in the Ohio program the items belonged. Nor did the PSE staff try to write an equal number of items for each developmental area. Instead, items were written to match actual classroom activities and apparent student learning. PSE believed that the number of items the staff wrote in each developmental area would reflect actual local program emphases. The distribution of items by developmental area would present a picture of the program as it was currently being implemented and could be used by the Division in diagnosing program strengths and weaknesses.

In order to assign items to developmental areas, PSE asked each local Program Director (or one of the local program staff members) in each of the four developmental sites to categorize each item by its

*Table 5 on page 38 lists the seven "developmental areas" into which the Ohio curriculum is divided.

developmental area after it was pilot tested. Every item was assigned a developmental area by a staff member from each of the four sites, yielding four opinions about each item.

The local program staffs did not always agree. This was partly because many items appeared to contain content relating to more than one developmental area. This was particularly true of affective items. Such items can usually be categorized as belonging in the "self" developmental area yet they also contain content from one or more of the other areas. That is, such items typically measure attitude toward one's self in relation to the world of work, employment and work adjustment skills, and so on. Wherever there was a lack of unanimity among the four local staff members categorizing the items--and unanimity was rare--PSE assigned the item to the category selected by the majority. If there were no majority, the PSE staff categorized the item on the basis of its knowledge of the State Curriculum Guides. (See Table 6, page 38.)

Conducting the Pilot Tests

All items derived from these procedures were pilot tested in four sites. Following is a description of the pilot testing process.

Pilot I. PSE administered its first pilot test--Pilot I-- on January 10 in Site 1. All the items which the staff had developed earlier in the week for grades 3, 6, 8, and 10 were typed and mimeographed by the Site 1 staff and administered to approximately 100 program students and 100 non-program students at each grade level. Several Site 1 Coordinators selected the program schools and the classrooms within them. The Site 1 Program Director used his knowledge of the

makeup of the local student population to choose comparable non-program schools and had the principals in those schools select the non-program classrooms. The PSE staff and the Site 1 Coordinators administered the tests and the PSE staff scored them.

One month later--February 3-7--PSE conducted Pilot 1 in the three remaining locations: Site 2, Site 3, and Site 4. The procedures duplicated those for Site 1 except that the PSE staff did not join the local Coordinators in administering the tests. A total of 100 program and 100 comparable non-program students in grades 3, 6, 8, and 10 took the tests at each of the three sites.

Pilot II. Upon returning from Item Writing II in Site 1, PSE printed and shipped pilot tests containing those items to all four sites. Again, the tests were administered to 100 program and 100 non-program students in each of grades 3, 6, 8, and 10 in each of the four sites. Program and non-program schools were once again selected by the local Program Directors and their Coordinators, duplicating the pattern used earlier.

More items were produced during Item Writing II than during Item Writing I because 1) PSE expanded the size of the writing team, 2) items were not pilot tested on site during the visit so that the staff could devote all its time to item writing, and 3) the staff was becoming more proficient in its item-writing skills. The result was a larger body of items than could be administered in a single class period.

In order to administer all the items, PSE produced two separate tests for each grade level, a Day 1 test and a Day 2 test, each requiring a full class period on two separate days.

Pilot III. All items that resulted from the Site 3 visit were included in Pilot III. The instruments were administered in all four sites. Again, the sample was 100 program and 100 comparable non-program students at each of the four grade levels.

Piloting Items at More Than One Grade Level. PSE assumed that items which were created as the result of an activity conducted at one grade level had the potential of discriminating at other grade levels as well. For example, an item which discriminated at the 3rd grade level might, if program knowledge were cumulative, discriminate at the 6th grade level as well. Following this assumption, PSE often piloted identical items at both the 3rd and 6th grades or at both the 8th and 10th grades. Occasionally an item was piloted at three or four grade levels if it seemed promising. In such cases, PSE adjusted the vocabulary of the item to accommodate to the reading level of students in the grades being tested.

Using Pilot Test Results to Rank Items by Discriminating Power

A "t" test was applied separately to the results for each item in each grade in each city to determine if there were a significant difference between program and non-program student responses. (Total test scores were not computed for individual students and mean test scores were not computed for classes or schools.)

Each item had the potential to 1) discriminate in favor of the program group; 2) be neutral (not discriminate between the two groups); or 3) discriminate in favor of the non-program group.

The items were, of course, intended to discriminate in favor of the program group and accordingly were written on the basis of information gathered in the field about current activities in program classrooms. Nevertheless, not all items worked as intended.

Each item was assigned a letter code based on the way it discriminated in the four pilot testing sites. For example, any cognitive item which discriminated in favor of the program group in all four pilot testing sites was assigned the letter A; any cognitive item which discriminated in favor of the program group in three out of the four sites and was neutral in the fourth, was assigned the letter B, and so on. This letter was the item's "discriminating power code". After an item was piloted, it was assigned its code and added to the item pool according to its ability to discriminate in favor of the program group.

Ranking Cognitive Items. Using a "+" to represent discriminating in favor of the program group, a "-" to represent discriminating in favor of the non-program group, and a "0" to represent neutrality, Table 1 presents in rank order from most desirable (most discriminating in favor of the program group) to least desirable (least discriminating in favor of the program group) the cognitive item discriminating power codes.

Ranking Affective Items. All affective items were written as agree/-disagree statements. While each cognitive item had a correct response, and the number of correct responses per program and non-program group was used to determine whether the item discriminated in favor of the program, many of the affective items the PSE staff wrote did not have a "correct" answer. Some affective items reflected attitudes which the program developers favored, but other items reflected self-concepts for which there is

Table 1

Basis for Assigning Discriminating Power Code to Cognitive Items

Pilot Tests

Pilot Site	Pilot Site	Pilot Site	Pilot Site	Discriminating Power Code
+	+	+	+	A
+	+	+	0	B
+	+	0	0	C
+	0	0	0	D
0	0	0	0	E
+	+	+	-	F
+	+	0	-	G
+	+	-	-	H
+	0	0	-	I
+	-	-	-	J
+	0	-	-	K
0	0	0	-	L
0	0	-	-	M
0	-	-	-	N
-	-	-	-	O

+ = Item discriminated in favor of program group

- = Item discriminated in favor of non-program group

0 = Item did not discriminate between program and non-program group

no correct response. Therefore, PSE ranked the affective items solely on their ability to distinguish between program and non-program groups. For example, an item with which a significant number of program students agreed in four out of four sites was assigned a discriminating power of AA; likewise, an item with which a significant number of non-program students agreed on in four out of four sites was also assigned an AA code. The more strongly an item moved in any given direction--more program students in more sites agreed or more non-program students in more sites agreed--the higher the ranking of the item.

Using a "+" to represent a significant number of program students agreeing with a statement, a "-" to represent a significant number of non-program students agreeing, and a "0" to represent neutrality, Table 2 presents in rank order from most discriminating to least discriminating the affective item discriminating power codes.

Selecting Items for the Final Test

In compiling the final test for each grade level, PSE planned to select the 40 cognitive items and the 20 affective items which discriminated most clearly in favor of program students.

The record kept on each item showed:

1. The item scores for program and non-program students.
2. The number of developmental sites in which the item discriminated between program and non-program students and its corresponding discriminating power code,
3. The developmental area(s) as assigned by the four pilot test site staffs.

Table 2

Basis for Assigning Discriminating Power Code to Affective Items*

Pilot Tests

Pilot Site	Pilot Site	Pilot Site	Pilot Site	Discriminating Power Code
+	+	+	+	AA*
-	-	-	-	
+	+	+	0	BB
-	-	-	0	
+	+	0	0	CC
-	-	0	0	
+	0	0	0	DD
-	0	0	0	
0	0	0	0	EE
+	-	0	0	FF
+	+	+	-	GG
-	-	-	+	
+	+	-	0	HH
-	-	+	0	
+	+	-	-	II

+ = Significant no. of program students agreed with item

- = Significant no. of non-program students agreed with item

0 = Item did not discriminate between program and non-program groups

*The content of an affective item determines whether it is preferable for any given item to have a significant number of program students agree or a significant number of non-program students agree.

PSE assembled the items into four bundles -- one for each grade level -- with the items ranked from the most discriminating (A for cognitive and AA for affective) to the least discriminating (Q for cognitive and II for affective). Some items appeared in more than one bundle, having been piloted at more than one grade level. An item's discriminating power usually differed at each grade level.

Selecting Cognitive Items. PSE took the top-ranked 40 items from each grade level bundle for the cognitive section of the test. If there were less than 40 A items, PSE used the B items. If there were still less than 40 items using all A and B items, PSE used the C items, and so on. When only some of the items in the final category were needed to reach the total of 40, PSE assembled data on each item in that category for all four sites combined and computed a "t" test for the combined data. (The original "t" tests were computed for each site separately, as already explained.) The results were used to rank items within the final category and enough items were then chosen in rank order of discriminating power to complete the required total of 40.

Selection was made without regard to item content or developmental area. PSE selected items solely on their ability to discriminate between program and non-program students.

Selecting Affective Items. Before selecting the top-ranked 20 affective items from each grade level bundle, PSE reviewed all the affective items and eliminated those which it felt would not yield helpful information to the program developers. Most of these were items in which a student was asked

to express personal preferences such as "I'd rather work alone than in a group," or "I'd rather work out-of-doors than in-doors." PSE took the position that the Ohio Career Development Program, while interested in helping students discover their own personal preferences, did not care what they preferred and thus would not be assisted by having data on the differences between the personal preferences of program and non-program students.

PSE had created affective items about personal preferences because its item writers have observed lessons or activities in which students discussed their likes and dislikes. Some of the resulting items did discriminate between program and non-program groups in the various pilot tests. However, PSE believed that items dealing with other topics reflected the objectives of the program developers and were equally representative of what program students were discussing and thinking about in school.

PSE retained the affective items which allowed students to express attitudes which the Career Development Program was interested in encouraging. Using this criterion, PSE retained items such as, "There are jobs I will like doing when I grow up," which reflects program values, but eliminated items such as, "I like to do things in groups," which is a matter of personal preference.

From those affective items retained, PSE selected 20 for the final tests, using the same procedures it had used for selecting the cognitive items for the final tests.

Scoring Affective Items

Affective items had been selected for the final tests on 1) their discriminating power (their ability to distinguish between program and non-program groups) and 2) their content (attitudes or feelings which PSE believed the program wanted to encourage). PSE presented to the Division the items it had selected and asked the Career Development staff to provide PSE with the answer to each item that best reflected program goals. The Division staff's answers to each affective item were later used to score the affective sections of the final tests.

Administering the Final Tests

Once the items for the final tests were selected, PSE began preparing for the final test administration in Akron, Cincinnati, Cleveland, East Muskingum, Mansfield, South-Western City, and Springfield.

Sample Size for the Final Tests. Classrooms rather than individual students were drawn in the sample for the final tests. Entire classrooms were tested, with all students taking part. In grades 8 and 10, either English or Social Studies classrooms were used.

PSE instructed the local Program Directors to test the following numbers of classrooms in each of the four sites:

1. 15 classrooms for each of the grades 3, 6, 8, and 10.
If fewer than 15 classrooms were available, all classrooms.
2. 5 non-program classrooms for each grade level. If fewer than 5 were available, all classrooms.

The oversampling of program classrooms was intended to allow later analyses of specific program effects.

Selecting Schools. Each school in the program was included in the

Population from which classroom teachers were drawn in administering the final tests.

Non-program schools were selected according to the following procedure. PSE asked the local Program Directors in each city to collect the following information for each school in the program where such information was available:

1. Mean reading scores
2. Mean I. Q. scores
3. Grade point averages
4. Any socio-economic indicators on file

For each of the four grade levels, the Program Director was asked to identify non-program schools whose students' scores most closely matched those of the program schools. Whenever possible, an equal number of non-program and program schools were selected.

When the test results were analyzed later, it became apparent that the local Program Directors had not been able to select entirely comparable schools. Measures such as the proportions of black/white/other students differed between the groups of schools, with the program schools having relatively more blacks and the non-program schools having relatively more whites. Such differences reflected local decisions about where to offer career education. Since the state subsidy had to be concentrated in a single high school and its feeder junior high and elementary schools, a local district which placed the program in a disproportionately black area sometimes had to draw non-program students from the remaining disproportionately white areas. The effect this had on the results is explained in Findings, beginning on page 58.

Selecting Classroom Teachers. Each Program Director was sent a list of 15 random numbers. They were asked to use a list of program teachers' names which were already ordered in some numbered fashion, and to select the names of the teachers whose numbers corresponded to the random numbers. Each Director was also sent a list of 5 random numbers to use with a similar list of non-program teachers' names.

Administering the Tests. PSE shipped the final tests to the seven Program Directors, who distributed them to the program and non-program schools. In some sites, the teachers administered the tests to both program and non-program students; in others, the Career Development Coordinators in each program building administered the tests to program students in their own buildings and to non-program students in other buildings.

Table 3 displays how many tests were administered to program and non-program students at each grade level in the seven sites.

Analyzing Data From the Final Tests. Data for cognitive and affective tests were analyzed by:

1. Program and non-program groups
2. Grade
3. City
4. State (seven cities combined)
5. Student background information (male/female and black/white other) and (at the 10th grade) type of curriculum being studied
6. Seven developmental areas

PSE generated item scores and total scores. The results are discussed in the Findings chapter.

Table 3

Administration

Final tests

Site	Number of Students									
	Grade 3		Grade 6		Grade 8		Grade 10		TOTAL	
	P*	NP**	P	NP	P	NP	P	NP	P	NP
Akron	368	131	426	129	375	114	341	132	1510	506
Cincinnati	359	137	342	138	260	91	242	96	1203	462
Cleveland	381	143	427	140	393	85	230	117	1431	485
East Muskingum	150	128	151	122	194	170	142	163	637	583
Mansfield	266	95	302	88	266	86	241	117	1075	386
Springfield	332	123	385	105	345	129	265	106	1327	463
South-Western City	385	177	391	203	444	208	300	182	1520	770
TOTAL	2241	934	2424	925	2277	883	1761	913	8703	3655

*p = Program

**NP = Non-program

OUTCOMES OF THE TEST DEVELOPMENT PROCEDURE

The test development procedure described previously produced a set of pilot test instruments consisting of all items written and a set of final test instruments consisting of those items which discriminated most favorably toward the program. Both sets of tests are described below.

The Pilot Tests

PSE wrote or selected 439 unique cognitive items and 161 unique affective items for pilot testing. (In addition to the items completed, PSE drafted or considered hundreds more.) A total of 208 of the cognitive items and 127 of the affective items were pilot tested at more than one grade level. The total count (unique items plus duplicated items) pilot tested for all four instruments combined was 683 cognitive items and 383 affective items. Table 4 shows the numbers of items pilot tested at each grade level.

Table 4

Number of Items Developed Pilot Tests

Grade	Cognitive items	Affective items
Grade 3	133	76
Grade 6	170	96
Grade 8	187	106
Grade 10	193	105
Total	683	383

Number of Pilot Test Items in Each Developmental Area

As explained earlier, each item developed and pilot tested was later assigned to one of the seven developmental areas in the Ohio Career Developmental Program by having the PSE staff collect and consolidate the opinions of local personnel in the four developmental sites as to the proper classification.

Cognitive Items. The numbers of cognitive items classified in each developmental area at each grade level are displayed in Table 6.

If the percent of items falling into each developmental area reflects present program cognitive emphases, then the developmental areas are being stressed in the following order at each of the four grade levels. These percents are shown in Table 5.

Table 5

Percent of Cognitive Items in Each Developmental Area Pilot Tests

Grade 3	% of Items Developed	Grade 6	% of Items Developed	Grade 8	% of Items Developed	Grade 10	% of Items Developed
WOW	27	WOW	30	WOW	37	WOW	26
I & E	18	I & E	15	ECON	17	ECON	21
ECON	13	ECON	13	EMP & WAS	13	ED & TR	17
ED & TR	13	ED & TR	13	I & E	11	EMP & WAS	11
SELF	12	SELF	12	SELF	9	SELF	10
DM	9	DM	9	ED & TR	7	I & E	8
EMP & WAS	8	EMP & WAS	8	DM	6	DM	7

Table 6

Number of Cognitive Items Developed

Pilot Tests

Developmental Area	Grade 3		Grade 6		Grade 8		Grade 10		Total	
	No.	%*	No.	%	No.	%	No.	%	No.	%**
Individual and Environment	24	18	25	15	21	11	15	8	85	13
World of Work	36	27	52	30	70	37	51	26	209	31
Economics	17	13	22	13	31	17	40	21	110	16
Education and Training	17	13	22	13	13	7	32	17	84	12
Employability and Work Adjustment Skills	11	8	14	8	24	13	22	11	71	10
Self	16	12	20	12	16	9	19	10	71	10
Decision Making	12	9	15	9	12	6	14	7	53	8
TOTAL	133	100	170	100	187	100	193	100	683	100

*Percent of items generated for each developmental area within a grade level.

**Percent of items generated for each developmental area in all grade levels combined.

Except for the very heavy emphasis on the World of Work, both the order of emphasis and the differences in the amount of emphasis (ranging from a low of 6 percent on Decision Making at grade 8 to a high of 21 percent on Individual and Environment in grade 10) is consistent with the goals of the Ohio Career Development Program. PSE is aware that as of 1975, the Division is encouraging local Program Directors to modify their efforts so as to focus less on the World of Work and more on the six other areas. As a consequence, the percent of items written to reflect the World of Work at each grade level should decrease when this item development process is repeated in coming years.

Affective Items. In the affective domain, the number of items PSE developed and pilot tested in each developmental area at each grade level is displayed in Table 7.

Ranking the developmental areas in order of apparent present program affective emphases, as suggested by the number of items developed in each developmental area, indicates that the developmental areas are being stressed in the order shown in Table 8.

Table 7

Number of Affective Items DevelopedPilot Tests

Developmental Area	Grade 3		Grade 6		Grade 8		Grade 10		Total	
	No.	%*	No.	%	No.	%	No.	%	No.	%**
Individual and Environment	6	8	7	8	8	8	9	9	30	8
World of Work	12	16	20	21	22	21	19	18	73	19
Economics	6	8	5	5	6	6	6	6	23	6
Education and Training	7	9	12	12	13	12	16	15	48	13
Employability and Work Adjustment Skills	2	3	6	6	12	11	11	10	31	8
Self	36	47	38	40	32	30	29	28	135	35
Decision Making	7	9	8	8	13	12	15	14	43	11
TOTAL	76	100	96	100	106	100	105	100	383	100

*Percent of items generated for each developmental area within a grade level.

**Percent of items generated for each developmental area in all grade levels combined.

Table 8

Percent of Affective Items in Each Developmental Area
Pilot Tests

Grade 3	% of Items Developed	Grade 6	% of Items Developed	Grade 8	% of Items Developed	Grade 10	% of Items Developed
SELF	47	SELF	40	SELF	30	SELF	28
WOW	16	WOW	21	WOW	21	WOW	18
DM	9	ED & TR	12	DM	12	ED & TR	15
ED & TR	9	DM	8	ED & TR	12	DM	14
I & E	8	I & E	8	EMP & WAS	11	EMP & WAS	10
ECON	8	EMP & WAS	6	I & E	8	I & E	9
EMP & WAS	3	ECON	5	ECON	6	ECON	6

As explained earlier, the local Program Directors and their staffs did not always agree on whether an item fell into one developmental area or another. This was particularly true of the affective items. If PSE had solicited more (or fewer) opinions when categorizing the items, Table 8 might look somewhat different. Nevertheless, the strong emphasis on Self is quite consistent with the affective goals of the Program.

Not surprisingly, the World of Work claimed the second largest number of items at all grade levels. However, as the Division succeeds in leading local personnel to decrease their attention to the cognitive domain of the World of Work in local classrooms, so too will they probably decrease their attention to the affective domain of the World of Work. If so, test items developed with the same procedure in future years should show a shift of attention into other developmental areas.

Discriminating Power of Pilot Test Items

Each item developed could, at one extreme, discriminate in favor of the program group in four cities, discriminate against the program group in four cities, or fall at some point between. (Items used at more than one grade level could work differently at each grade and thus were scored, analyzed, and reported as if they were a separate item at each grade.)

The Discriminating Power Code was explained earlier. See Table 1, page 27, for cognitive items and Table 2, page 29, for affective items.

Direction of Discrimination. In which direction did the items discriminate: for the program or against the program?

Approximately half of all the cognitive items developed discriminated in favor of the program in one or more pilot testing sites. About 25 percent did not discriminate in either direction in any site and about 25 percent discriminated against the program in one or more sites. In other words, of the 683 cognitive items PSE tested (including duplicates used at more than one grade level) two out of four favored the program students, one out of four was neutral, and one out of four favored the non-program students.

Approximately half of all the affective items discriminated in favor of the program in one or more pilot testing sites. About 35 percent did not discriminate in either direction in any site, and about 15 percent discriminated for the program in some sites and against it in others in what seemed to be a random manner. (Upon closer inspection, most of the latter group consisted of what PSE came to call "personal preference" items. They were subsequently eliminated, for reasons given earlier.)

In short, of the 383 affective items PSE tested (including duplicate items used at more than one grade level) three out of six discriminated in favor of the program, two were neutral, and one behaved in a mixed, unpredictable fashion.

The numbers of items which were tested compared to the numbers of items which discriminated are displayed in Tables 9, 10, and 11. An examination of these tables shows that cognitive items were most likely to discriminate in favor of the program at grade 3 (68 percent of the items), less likely to do so at grades 6 and 8 (50 percent and 60 percent, respectively), and least likely to do so at grade 10 (32 percent). To express the same idea somewhat differently, it was easier for the PSE staff to identify cognitive learning in program classes at grade 3, more difficult at grades 6 and 8, and most difficult at grade 10. This finding is squarely in keeping with PSE's field observations carried on in program classrooms during the past three years.

The same is not true for student learning in affective areas. An examination of Tables 9, 10, and 11 indicates that about half of the affective items written discriminated in favor of the program at each of the four grade levels--with only minor differences between grades. To be exact, affective learning was most evident in grades 6 and 8, somewhat less evident in grade 10 and least evident in grade 3--according to these findings. What appears to be happening is that grade 3 program teachers pay a lot more attention to the cognitive aspects of careers than non-program teachers do, but pay only somewhat more attention to the affective aspects of careers than non-program teachers do. But the

Table 9

Number of Items Developed Compared to Number of Items
Which Discriminated in FAVOR of the Program in One or More Pilot Testing Sites*

Pilot Tests

Grade Level	Cognitive Items			Affective Items		
	Number Developed	No. Discriminating in 1 or more Sites	% Discriminating in 1 or more Sites	Number Developed	No. Discriminating in 1 or more Sites	% Discriminating in 1 or more Sites
3	133	91	68	76	32	42
6	170	86	50	96	54	56
8	187	112	60	106	62	58
10	193	62	32	105	53	50
Total	683	351	51	383	201	52

*Out of a total of 4 pilot testing sites

Table 10

Number of Items Developed Compared to Number of Items
Which Discriminated AGAINST the Program in One ore More Pilot Testing Sites*

Pilot Tests

Grade Level	Cognitive Items			Affective Items		
	Number Developed	No. Dis-criminating in 1 or more Sites	% Dis-criminating in 1 or more Sites	Number Developed	No. Dis-criminating in 1 or more Sites	% Dis-criminating in 1 or more Sites
3	133	15	11	76	16	21
6	170	45	26	96	10	10
8	187	30	16	106	18	17
10	193	79	41	105	10	10
Total	683	169	24	383	54	14

*Out of a total of 4 pilot testing sites

Table 11

Number of Items Developed Compared to Number of Items
Which DID NOT DISCRIMINATE in Any Pilot Testing Site*

Pilot Tests

Grade Level	Cognitive Items			Affective Items		
	Number Developed	No. Not Discriminating in Any Site	% Not Discriminating in Any Site	Number Developed	No. Not Discriminating in Any Site	% Not Discriminating in Any Site
3	133	27	20	76	28	37
6	170	39	23	96	32	33
8	187	45	24	106	26	25
10	193	52	27	105	42	40
Total	683	163	23	383	128	33

*Out of a total of 4 pilot testing sites

reverse seems to be true at grade 10. That is, grade 10 program teachers pay a lot more attention to the affective aspects of careers than non-program teachers do but pay only somewhat more attention to the cognitive aspects of careers than the non-program teachers do. (In grades 6 and 8, program teachers pay somewhat more attention to both the cognitive and affective aspects of careers than non-program teachers do.)

That is not particularly surprising, when looked at this way: most grade 3 teachers attend to the affective aspects of careers and most grade 10 teachers attend to the cognitive aspects of careers--even though they are not in the program. It is in the areas they tend to neglect--cognitive aspects in grade 3 and affective aspects in grade 10--that the program teachers show their superiority. Similarly, if grade 6 and grade 8 teachers already give relatively equal attention to the cognitive and affective aspects of careers, program teachers are able to show some superiority to them in both respects.

The Final Tests

Two elements of the final tests are of special interest:

1. The number of items falling into each developmental area at each grade level,
2. The discriminating power of the items.

Together, these two elements present a picture of the Ohio Career Development Program at its strongest: the areas of emphasis and the most discriminating items within those areas.

Number of Final Test Items in Each Developmental Area

PSE took the 60 best items developed at each grade level and included them in its final tests. (The term "best" means discriminating most strongly in favor of program students.) No effort had been made to develop and none was made to select an equal number of items in each of the seven developmental areas. Indeed, PSE and the Division were prepared for the possibility that there might not be any items in some of the developmental areas. However, the PSE staff did observe activity in all seven areas and wrote items accordingly.

It is a credit to the Ohio Program that the set of 40 best cognitive items at each grade level contained questions in all seven developmental areas and the set of 20 best affective items contained questions in all seven developmental areas in three out of the four grade levels.

Cognitive Items. The number of cognitive items in each developmental area ranged from a high of 14 World of Work items in the grade 8 test to a low of 2 Decision Making items in the grade 8 test. Assuming that the number of items in each developmental area is indicative of current program emphases, it is useful to look at the number of items per area. Table 12 ranks the seven areas according to the number of items in each. Table 13 presents the same data without ranking the seven areas. The distribution of discriminating final test items is similar to the distribution of all items developed and pilot tested, as shown in Table 5 and Table 6 earlier (see pages 37 and 38). Thus the quantity as well as the quality of items reveals a similar pattern of attention to the seven developmental areas across the four grade levels.

Table 12

Number of Cognitive Items in Each Developmental Area

Final Tests

Grade 3	No. of Items Selected	Grade 6	No. of Items Selected	Grade 8	No. of Items Selected	Grade 10	No. of Items Selected
I & E	10	WOW	10	WOW	14	WOW	8
WOW	8	ED & TR	8	I & E	9	ECON.	7
ED & TR	7	ECON.	6	EMP & WAS	6	SELF	6
SELF	5	SELF	5	ECON	3	ED & TR	5
ECON	4	DM	5	ED & TR	3	EMP & WAS	5
DM	3	I & E	3	SELF	3	DM	5
EMP & WAS	3	EMP & WAS	3	DM	2	I & E	4

An examination of Table 12 leads to the following general observations, each conditioned on the premise that the pattern displayed by these test items actually reveals local program emphases:

- The overall rankings are in keeping with the objectives of the program.
- The World of Work is emphasized in grades 3, 6, and 8 at the expense of other areas.
- The grade 10 program is best balanced across all areas.
- The attention given to the World of Work and to Individual and Environment does not decline as it should as the program progresses from grade 3 through grade 10.
- The attention given to Education and Training and to Employability and Work Adjustment Skills does not rise as it should from grade 3 through grade 10.

Table 13

Number of Cognitive Items in Each Developmental Area*

Final Tests

Developmental Area	Grade 3	Grade 6	Grade 8	Grade 10	Total
Individual and Environment	10	3	9	4	26
World of Work	8	10	14	8	40
Economics	4	6	3	7	20
Education and Training	7	8	3	5	23
Employability and Work Adjustment Skills	3	3	6	5	17
Self	5	5	3	6	19
Decision Making	3	5	2	5	15
TOTAL	40	40	40	40	160

*Items categorized by 4 pilot testing site Program Directors and their staffs.

- The areas of Self and Economics received fairly balanced attention from grade 3 through grade 10.
- Decision Making receives uneven--and perhaps inadequate--attention from grade 3 through grade 6.

Affective Items. Many affective items written by the PSE staff were classified by local Program Directors and their staffs as falling into the area of self--as might have been expected. The distribution of test items over the seven developmental area is shown in Table 14, with the areas ranked according to the number of items they contain. Table 15 displays the same information, without ranking the areas.

Table 14

Number of Affective Items in Each Developmental Area

Final Tests

Grade 3	No. of Items Selected	Grade 6	No. of Items Selected	Grade 8	No. of Items Selected	Grade 10	No. of Items Selected
SELF	8	SELF	9	SELF	7	SELF	5
WOW	7	DM	3	WOW	3	I & E	4
I & E	3	I & E	2	ED & TR	3	ED & TR	4
ECON	1	WOW	2	DM	3	DM	3
ED & TR	1	ED & TR	2	EMP & WAS	2	EMP & WAS	2
EMP & WAS	0	ECON	1	I & E	1	WOW	1
DM	0	EMP & WAS	1	ECON	1	ECON	1

The more consistent pattern of area emphasis exhibited by the distribution of items in Table 7 and Table 8 (see pages 40 and 41) presents a truer picture than the somewhat erratic display in Table 14, given the scarcity of strongly discriminating affective items available to fill Table 14.

Table 15

Number of Affective Items in Each Developmental Area*

Final Tests

Developmental Area	Grade 3	Grade 6	Grade 8	Grade 10	Total
Individual and Environment	3	2	1	4	10
World of Work	7	2	3	1	13
Economics	1	1	1	1	4
Education and Training	1	2	3	4	10
Employability and Work Adjustment Skills	0	1	2	2	5
Self	8	9	7	5	29
Decision Making	0	3	3	3	9
TOTAL	20	20	20	20	80

*Items categorized by 4 pilot testing site Program Directors and their staffs.

There are several possible interpretations of the data in Table 14. The most significant is the question they raise about whether the seven areas represent a satisfactory classification scheme for affective items. One alternative scheme for thinking about affective career learning is to envision that decisions must be made concerning the self in relation to the other five developmental areas: decisions about the self in relation to the world of work, in relation to the amount of education and training one should obtain, and so on. Such a structure would call for constructing affective test items all of which would be classified under self and decision making and under one (or more) of the other five areas as well.

Another interpretation of the data in Table 14 is that state and local Program leaders have not yet trained teachers to carry on a full array of affective instruction across all seven areas.

Yet another interpretation is that affective behaviors are not required--or at least not required equally--in each of the seven areas.

Any one of these interpretations is consonant with the data. Whether one of these interpretations--or some other one--is in fact correct, it does seem clear that state and local officials in Ohio need to give further thought to the affective aspects of career development in the seven areas.

Discriminating Power of the Items Selected for the Final Tests

Table 16 displays the discriminating power of the cognitive items selected for the final tests while Table 17 displays parallel information for affective items. As explained earlier, all of the items in the most discriminating categories (A for cognitive and AA for affective) were

Table 16

Discriminating Power of Cognitive Items Selected

Final Tests

Discriminating Power Code*	Number of Items				
	Grade 3	Grade 6	Grade 8	Grade 10	TOTAL
A	3	0	0	0	3
B	16	6	16	0	38
C	21	23	24	9	77
D	0	11	0	31	42
TOTAL	40	40	40	40	160

*The explanation of the discriminating power code appears earlier.

Table 17

Discriminating Power of Affective Items Selected

Final Tests

Discriminating Power Code*	Number of Items				
	Grade 3	Grade 6	Grade 8	Grade 10	TOTAL
AA	0	2	0	0	2
BB	0	0	1	0	1
CC	4	2	4	4	15
DD	10	16	12	11	49
EE	1	0	0	1	2
FF	2	0	2	2	6
GG	1	0	0	0	1
HH	2	0	1	1	4
TOTAL	20	20	20	20	80

*The explanation of the discriminating power code appears earlier.

used before selecting items from the lower, less discriminating categories.

Table 16 shows that PSE was able to obtain the required 40 cognitive test items for grade 3 and grade 8 without going beyond the first three categories; that is, those two tests were compiled entirely from items which discriminated in favor of the program in four sites, or three sites, or (at worst) 2 sites.

In contrast, 11 of the 40 items (28 percent) needed for the grade 6 test had to be chosen from those that discriminated in only 1 site.

And, still worse, 31 of the 40 items needed for the grade 10 test (78 percent) had to be drawn from items that discriminated in only 1 site.

This pattern of discriminating power made it possible to predict, even before the final tests were administered, that the ultimate findings would be more favorable for grade 3 than for grade 10, with grades 6 and 8 falling between the two extremes.

By contrasting Table 16 with Table 17, one can see how much more discriminating were the cognitive items than the affective items that went into the final tests. While Table 16 shows the cognitive tests could be completed with items drawn from the top four categories, Table 17 shows it was necessary to go into the next four categories to obtain the 20 affective items needed at grade 3, grade 8, and grade 10.

There are few differences in the pattern at the four grade levels, except that the grade 6 test was comprised of somewhat more discriminating items.

Taken as a whole, Table 17 shows that the PSE staff had difficulty both in identifying affective learning in program classrooms and in writing test items on which program students could demonstrate their superiority.

The second explanation seems to be more powerful than the first. That is, the PSE staff did find evidence of affective learning in program classrooms; the real difficulty came in writing test items that would demonstrate the superiority of that learning to the affective learning simultaneously taking place in non-program classrooms

Affective learning is widely recognized as being very difficult to measure. While it may be that in the spring of 1975 there are few real differences in affective learning between program and non-program classrooms in the Ohio program, it is possible that better instruments--instruments taking longer to develop and longer to administer--would show that the program students are superior.

FINDINGS

The study had two major outcomes:

1. A set of test results for grades 3, 6, 8, and 10 in seven cities.
2. A set of test instruments for grades 3, 6, 8, and 10 with potential use statewide.

The two outcomes are presented in separate sections on the following pages.

All of the conclusions presented are supported by data in the tables. Main points are supported by short tables placed in the text. Subordinate points are supported by the longer numbered tables. All numbered tables for the Findings section appear at the end of this volume.

TEST RESULTS

The results of the final tests are reported in two ways:

1. For the state as a whole (the state is represented by the seven cities combined).
2. For the seven individual cities.

Because what is reported below as state results is nothing more or less than results for the seven cities combined, whatever is said about the state is true--on the average--for the individual cities. However, the cities differ sufficiently among themselves that while a conclusion about average ("statewide") performance applies to the typical city, it does not apply to every individual city.

Overall Results for the State

Although the seven cities participating in the final tests differ enormously in total school population--ranging from Cleveland at one extreme to East Muskingum at the other--it is very important to remember that the final tests were administered to approximately the same number of program students in each of the seven cities and to approximately the same number of non-program students in each of the seven cities.* This means that East Muskingum contributed about as much weight to the statewide results as Cleveland. Thus good results in a large city can be largely offset by poor results in a small city, and vice versa.

The decision to give each city equal weight was based partly on the fact that the project sites within cities are approximately equal in student population (since each consists of a senior high school and its feeder junior high and elementary schools) and partly on the desire to apply both the results (and ultimately the instruments) to all kinds of cities statewide.

- Program students scored higher than non-program students in grades 3, 6, 8, and 10 on both cognitive and affective tests.
- The superiority of the program students was statistically highly significant on most tests.
 - It reached the .01 level on all tests except the grade 8 affective test (.05) and the grade 10 affective test (positive but less than .05).

*City 5 supplied only about half as many program students as the other cities. (See Table 18.)

- Program students showed their greatest superiority over non-program students in grade 3, next in grade 6, next in grade 8, and last in grade 10.
- Program students showed greater superiority on the cognitive tests than on the affective tests.

State Total Test Results*

Final Tests

Test	Grade	"t" value***
Cognitive	3	16.73**
Cognitive	6	12.81**
Cognitive	8	9.60**
Cognitive	10	5.56**
Affective	3	5.56**
Affective	6	4.40**
Affective	8	2.03*
Affective	10	1.04

*Significant at the .05 level.
 **Highly significant at the .01 level.
 *** "t" values are for mean differences between program and non-program students.

- The typical student--both program and non-program--found the cognitive tests relatively easy, although not as easy as the affective tests. (Data to support this statement and the related statements which follow appear in Tables 19).
 - o On the cognitive tests, the typical program student answered just over 60 percent of the items correctly while the typical non-program student answered just under 60 percent correctly.
- The typical student--both program and non program--found the affective tests even easier than the cognitive tests.

*See Table 19 for details.

- o On the affective tests, the typical program student answered about 70 percent of the items correctly while the typical non-program student answered just under 70 percent correctly.
- The fact that the typical non-program student answered about 60-70 percent of the items correctly indicates that the content of the program is not unique to program classrooms but is being taught in non-program classrooms as well. While the program students are learning more of it, the non-program students are learning much of it.
 - o One interpretation is that the program designers have not yet chosen unique or distinctive content.
 - o Another interpretation is that they have chosen distinctive content but have not yet fully communicated it to classroom teachers.
 - o Another interpretation is that the content is not meant to be distinctive but rather the opposite: that is, it deals with important ideas and skills already in the regular school curriculum but attempts to teach them better by relating them to careers.
 - o Another interpretation is that the program designers have chosen ideas that are central to career success in our society--so central that they are transmitted to all students by their families, by their friends, by non-school organizations, and by the mass media.
 - o The superiority of the program students can be attributed to the fact that their schools supplement and reinforce information coming to all students from school and non-school sources.
- Regardless of the fact that non-program students scored well, program students clearly scored even higher than non-program students at all grade levels on both cognitive and affective tests.

The results throw light on the four interpretations hypothesized in 1973 when the curriculum-based tests failed to show differences between program and non-program students (see pages 6-8):

Interpretation 1: No Career Education Is Taking Place. This interpretation can now be firmly rejected.

Interpretation 2: Other School Curricula Teach Career Concepts. This interpretation continues to seem quite reasonable. Some of the test items that non-program students answered correctly appeared to require school-taught knowledge.

Interpretation 3: Non-School Sources Teach Career Concepts. This interpretation continues to seem especially reasonable. Many of the questions that non-program students answered correctly required information they could have learned outside of school.

Interpretation 4: Program Students Cannot Fully Demonstrate Their Learning on the Tests. This interpretation can now be fully accepted. It is clear that the 1973 curriculum-based tests did not let the program students demonstrate their superiority as revealed by the 1975 field-based tests.

Overall Results for Seven Individual Cities

All of what was said about the state is true for the typical city--although not for every city--since the results for the state are simply a summary of the results for the cities. In the typical city, program students scored significantly higher than non-program students in grades 3, 6, 8, and 10--in that order--on both the cognitive and the affective tests. They showed greater superiority on the cognitive tests than on the affective tests. The typical student--non-program as well as program--found both tests relatively easy and answered about 60 percent of the cognitive items correctly and about 70 percent of the affective items correctly.

- In 28 cognitive test comparisons (4 grades x 7 cities), program students were superior in 21, the same in 3, and inferior in only 4. All differences were significant at the .01 level.
- In 28 affective test comparisons, program students were superior in 14 (12 at the .01 level), the same in 10, and inferior in only 4 (2 at the .01 level).

City Test Results Ranked by How Well
Tests Discriminated in
Favor of Program Students*

Final Tests

City Codes: Cities Listed in Rank Order	Grade 3		Grade 6		Grade 8		Grade 10	
	"t" value*		"t" value*		"t" value		"t" value	
	Cogni- tive Test	Affec- tive Test	Cogni- tive Test	Affec- tive Test	Cogni- tive Test	Affec- tive Test	Cogni- tive Test	Affec- tive Test
3	13.22**	2.37*	14.62**	4.84**	17.49**	6.31**	6.34**	1.59
7.	2.66**	2.89**	6.13**	3.05**	11.08**	6.06**	9.78**	.75
5	9.43**	5.76**	3.92**	2.71**	3.97**	.48	1.49	3.00**
1	10.49**	-.33	4.58**	3.19**	5.22**	-2.22**	6.93**	-1.32
6	.76	4.14**	-5.78**	-2.47*	11.34**	7.11**	4.76**	3.87**
2	7.22**	1.79	10.41**	2.36*	1.63	-1.10	-4.72**	1.79
4	5.79**	1.40	3.99**	.08	-11.05**	-3.58**	-6.57**	-2.74**
State Totals***	16.73**	5.56**	12.81**	4.40**	9.60**	2.03*	5.56**	1.04

*Significant at the .05 level.

**Significant at the .01 level.

***State Totals = data for seven cities combined.

- But there were clear differences among the seven cities, as shown in the table above.
 - o At the upper extreme represented by City 3, program students scored far above non-program students on the cognitive tests in grades 3, 6, 8, and 10 and scored well above them on the affective tests in grades 3, 6, and 8 as well as slightly above them in grade 10. At the lower extreme represented by City 4, program students were actually inferior overall to non-program students.

*See Table 20 for details.

- o Program students in some cities were actually far inferior to non-program students in other cities on both cognitive and affective tests.* For example, program students in City 2 and City 6 scored far below non-program students in City 3, City 4, and City 7 on both tests in all four grades. (However, these variations rather clearly reflect differences in student characteristics and may not indicate differences in program quality. More is said about this later when results are reported by student characteristics.)

What Do the City Rankings Signify?

- The city rankings do signify that:
 - o Program students in high-ranking cities know more about careers--as measured by the PSE tests--than students in low-ranking cities.
- But the city rankings do not signify that:
 - o These cities would rank in the same order on some other career test--particularly one written and tested in cities different from the ones in which the PSE tests were written and tested.
 - o These cities would rank in the same order if each city had been able to find fully comparable program and non-program students to test. City 2, for example, would have made a much better showing if the black/white/other proportions in program and non-program classes had been the same.
 - o Students in low-ranking cities are not learning anything valuable about careers. Indeed, it is actually possible that they are learning as much as or more than students in high-ranking cities, cannot show it on PSE tests, but could show it on other tests.

State Test Results for Students With Various Characteristics

The test results were analyzed to see whether they differed for different kinds of students. The major findings for students with different demographic characteristics (boys/girls and blacks/whites/other) and with different school-related abilities (English grades, mathematics grades, and socio-economic status as indicated by education of head of household and with different curricular interests (general/academic/vocational/other) are reported below.

*See Table 20.

Boys and Girls (see Table 21)

A study of the tables shows the following:

- In cognitive learning, the program works about as well for boys as for girls.
 - o Program boys and girls both scored higher than non-program boys and girls in grades 3, 6, 8, and 10 on the cognitive tests.
 - o All cognitive test differences were significant beyond the .01 level.
 - o Program boys and girls showed their greatest superiority over non-program boys and girls in grade 3, next in grade 6, next in grade 8, and last in grade 10 on the cognitive tests.
- In affective learning, on the other hand, the program works better for boys than for girls.
 - o Program boys scored higher than non-program boys in grades 3, 6, and 8, while program girls scored higher than non-program girls only in grade 3.
 - o The differences in affective learning were smaller than the differences in cognitive learning.

State Total Test Results for Boys and Girls

Final Tests

Test	Grade	"t" value***	
		Boys	Girls
Cognitive	3	12.43**	11.22**
Cognitive	6	8.14**	9.79**
Cognitive	8	5.15**	8.00**
Cognitive	10	5.69**	2.77**
Affective	3	4.37**	3.27**
Affective	6	4.46**	1.64
Affective	8	2.12*	0.18
Affective	10	1.04	0.50
*Significant at the .05 level. **Highly significant at the .01 level. ***"t" values are for mean differences between program and non-program students.			

- Girls start the program knowing as much about careers as boys. While both boys and girls gain information during the program, girls pull slightly ahead of boys in grades 6 and 8 but boys pull ahead to match them again by grade 10.
- Girls have slightly better attitudes toward careers than boys at all grade levels.

State Total Test Results

Final Tests

Percent of Items Answered Correctly by Program Students

Test	Grade	Boys	Girls
Cognitive	3	59	59
Cognitive	6	62	64
Cognitive	8	62	65
Cognitive	10	61	61
Affective	3	65	66
Affective	6	77	78
Affective	8	68	72
Affective	10	67	69

- Non-program girls do as well on the tests as program boys (see Table 21).
 - Evidently it takes a formal career education program to teach some boys the career information and attitudes that most girls learn elsewhere.

Blacks, Whites, and Others (see Table 22)

A study of the tables shows the following:

- In cognitive learning, the program works quite well for all three groups of program students: white, black, and other. All of them scored significantly higher than comparable non-program students in grades 3, 6, 8, and 10--generally in that order--on the cognitive tests.

*The discussion of change over time in this section grows out of comparing the status of program and non-program students at grades 3, 6, 8, and 10 in the spring of 1975. Since time-series data were not collected, the statements about changes are somewhat speculative.

- Whites start out ahead of blacks and others and then proceed to widen the gap just a bit, particularly the gap between themselves and blacks, particularly in grades 3 and 6.
- In affective learning, the program works quite well for all three groups of program students: white, black, and other. All of them scored significantly higher than comparable non-program students in 3 out of 4 grades tested (a different set of three for each group) on the affective tests.
 - o Whites start out ahead of blacks and others and then proceed to widen the gap just a bit between themselves and blacks as time passes. Other students, in contrast, tend to close the gap between themselves and whites as time passes.
 - o The differences in affective learning were smaller for all three groups than the differences in cognitive learning.

State Total Test Results
for Blacks/Whites/Others

Final Tests

Test	Grade	"t" value***		
		Blacks	Whites	Others
Cognitive	3	7.09**	21.73**	8.34**
Cognitive	6	7.00**	17.00**	6.50**
Cognitive	8	6.00**	14.96**	8.52**
Cognitive	10	5.41**	9.55**	3.03**
Affective	3	0.13	8.00**	3.18**
Affective	6	3.30**	5.33**	0.38
Affective	8	3.63**	2.45*	2.67*
Affective	10	3.07**	0.82	3.21**
*Significant at the .05 level. **Highly significant at the .01 level. ***"t" values are for mean differences between program and non-program students.				

The Ohio Career Development Program is not deliberately designed for students of any particular background, but is designed to serve them

all. The result is that all kinds of students improve in the program, with the initial gaps between them changing only slightly as a result of being in the program.

Grades and Socio-Economic Status

PSE attempted to collect intelligence test scores or achievement test scores from program and non-program schools in the seven cities. However, many schools were reluctant or unwilling to supply such information. Where the data were available and were supplied to us, they were rarely comparable from school to school and were even less comparable from city to city. As a substitute, we used other data.

Data supplied by non-program students in grade 8 and in grade 10 were analyzed to determine what correlation existed between the final grades the students had received in English and in mathematics last year, the education of the heads of the students' households, and scores on the cognitive tests and the affective tests. The grades as an indicator of school achievement and the education of household head as an indicator of socio-economic status were regarded as independent variables which might be related to the two test scores as dependent variables. If so, those relationships might explain differences between the performance of program and non-program students in various grades in various cities.

The analysis was made for "untreated" non-program students so as to distinguish the effect of background factors on test scores from the effect of the program on test scores. The results of the analysis are shown in the table on the following page.

Correlations Between Grades,
Socio-Economic Status, and Test Scores

Final Tests

Grade 8 and Grade 10 Non-Program Students

Measure	Measure			
	Education of Head of Household	English Grades	Math Grades	Cognitive Test
English Grades	.21			
Math Grades	.15	.29		
Cognitive Test	.27	.36	.16	
Affective Test	.15	.22	.11	.47

It is clear that while there is a positive relationship between both cognitive and affective test scores and grades received in English and mathematics as well as socio-economic status, those relationships are so low that the test scores cannot be explained by the other measures. (PSE employed multiple regression techniques to determine whether a combination of grades and socio-economic status might have a higher correlation with the cognitive and affective test scores. While the combination showed a slightly stronger relation than the individual measures, even the combination was too low to be worth pursuing.) In short, the differences in cognitive and affective test scores cannot be explained by differences in grades and socio-economic status, insofar as grades and status are accurately measured by what the grade 8 and grade 10 non-program students reported about themselves. More refined, more accurate measures might have shown a stronger relation.

The low correlation between cognitive test scores and affective test scores indicates that those tests are measuring different factors which

"overlap" each other by about 25%. Presumably what students know about careers affects how they feel and the way they feel affects what they learn about careers.

Grade 10: Academic, Vocational, General, and Other Curricula (See Table 23)

PSE asked students in grade 10 to indicate the kinds of courses they were taking, on the assumption that their answers might be related to their test scores.

The analysis of their answers showed clear differences among students' test scores according to the curriculum in which they were enrolled. But it showed few differences in the effect of the career development program according to their curricula.

- Both program and non-program students scored on both the cognitive tests and the affective tests according to their curricula, in the following rank order from high to low: academic, vocational, general, and other.
 - o Academic students--both program and non-program--far outstripped the three other groups. The differences between the three other groups themselves were relatively small.
 - o Being enrolled in an academic curriculum may be correlated with being white. If so, it is not clear whether the superiority of the academic students is traceable to their backgrounds, their curriculum, a combination of the two, or perhaps a third causative factor which happens to be correlated with being white and being in an academic curriculum.

State Total Test Results

Final Tests

Percent of Items Answered Correctly*

Test	Grade 10 Curriculum							
	Academic		Vocational		General		Other	
	Program	Non-Program	Program	Non-Program	Program	Non-Program	Program	Non-Program
Cognitive	69	70	58	56	56	55	54	51
Affective	72	73	67	66	65	64	63	62
*See Table 23 for details.								

- The actual impact of the program on student learning is not closely related to the curriculum in which the students say they are enrolled. Differences showed in only 2 out of 8 comparisons.
- In cognitive learning, the program appears to be most effective for students reporting that they are enrolled in a curriculum "other" than academic, vocational, or general; next best for students who say they are enrolled in a general curriculum; and evidently not at all effective for students who say they are enrolled in an academic or vocational curriculum.
 - o A tentative but inconclusive interpretation is that students in academic and vocational programs have already acquired sufficient information and attitudes about careers to have set their post-secondary school goals--some heading for college and some heading for immediate employment--while students in unclassified "general" or "other" curricula have not yet acquired such information and attitudes. It would logically follow that the second group would be more susceptible to influence by a career development program since it could supply the information and attitudes they lacked.
 - o The fact that academic students enrolled in the program do not do quite as well as academic students not enrolled in the program probably can be explained by the fact that the seven cities show a tendency to locate the program in senior high schools which need student learning strengthened. If the academic students in such high schools are not as strong scholastically as the academic students in non-program schools, this could explain why the non-program students score higher.

State Total Test Results

Final Tests

"t" values*

Test	Grade 10 Curriculum			
	Academic	Vocational	General	Other
Cognitive	-1.69	1.73	2.36*	4.34**
Affective	-1.82	0.88	1.12	0.12

*Significant at the .05 level.
**Highly significant at the .01 level.
***"t" values are for mean differences between program and non-program students.

City Test Results for Students With Various Characteristics

All of what was said about the state is true for the typical city.

"Girls in grades 3, 6, 8, and 10 have slightly better information and attitudes about careers than boys, especially in the middle grades, but since the program works a bit better for boys, the boys almost manage to close the gap by the time they reach grade 10. Different patterns of maturation for boys and girls may be at work, pushing the girls ahead during early adolescence but allowing the boys to overtake them by about age 15. Whites start out far ahead of blacks and others in grade 3 and, because the program works a bit better for whites, they widen the gap just a bit by the time they reach grade 10. Even so, the program works well for all three groups: blacks, whites, and others. Grade 10 academic students--both program and

non-program—score above students in other curricula on both the cognitive and the affective tests. And the academic students outside the program score above students inside the program, probably because the program is usually placed in senior high schools whose academic students need help.

- But there were clear differences among the seven cities. Not every city displayed the typical pattern.
 - In City 7 the program worked better for boys than for girls overall on both cognitive and affective tests. In City 1 it worked better for girls than for boys overall. And in City 3 it worked extremely well for both boys and girls overall.
 - In City 2 about 90 percent of the students are black while about 10 percent are white. If these proportions had been reversed, program students in City 2 would have scored much higher on both the cognitive tests and the affective tests. In City 6, the same pattern occurs but to a lesser degree. And in City 5 both program and non-program students are virtually 100 percent white. (These city figures may be contrasted with the state totals showing about 30 percent of the program students and about 15 percent of the non-program students to be black.)
 - In City 6 and City 7 the grade 10 academic program students outstripped their non-program counterparts. In City 2 and City 4, the reverse occurred. And in City 1 and City 2 and City 5 there was no difference.
 - In City 1 and City 6 the vocational students in the program scored far ahead of the vocational students outside the program. There was a slight but undramatic reversal of this pattern in City 2 and City 4. In the other cities, there was no difference.
 - City 4 separated itself from the majority when its program students enrolled in the general curriculum and the "other" curriculum scored lower than non-program students in those same curricula. But this was almost the sole exception. Other cities followed the pattern for the state as a whole with program students in general and "other" curricula getting better scores than non-program students in those curricula.

TEST INSTRUMENTS

The test development activities yielded several outcomes:

1. A set of four tests -- each with cognitive and affective subtests -- for grades 3, 6, 8, and 10.
2. A set of additional test items not used in the four tests but suitable for possible future use in developing new tests.
3. A group of extremely promising test development procedures that merit refinement through further use.

Each of the outcomes is discussed below.

The Four Tests

Each of the four cognitive tests consists of 40 multiple-choice paper-and-pencil items classified into the seven developmental areas of the Ohio Career Development Program. Each of the affective tests consists of 20 multiple-choice paper-and-pencil items similarly classified.

The individual items in the tests differed sharply in their ability to discriminate in favor of program students; some items did it very well; others did not do it at all; and a few discriminated against program students.

The cognitive tests discriminated in favor of program students more clearly than the affective tests. Both discriminated in favor of program students at the lower grades more clearly than at the upper grades. (In fact, the grade 10 affective test did not discriminate significantly in favor of program students.) Both patterns are shown in the table below.

State Total Test Results Ranked by
Ability of Test to Discriminate in
Favor of Program Students

Test	Grade	"t" value***
Cognitive	3	16.73**
Cognitive	6	12.81**
Cognitive	8	9.60**
Cognitive	10	5.56**
Affective	3	5.56**
Affective	6	4.40**
Affective	8	2.03*
Affective	10	1.04

*Significant at the .05 level.
 **Highly significant at the .01 level.
 *** "t" values are for mean differences between program and non-program students.

The ability of the cognitive tests and the affective tests to discriminate in favor of program students differed sharply among the seven cities at the four grade levels. The grade 3 cognitive test did it in six of the seven cities; the grade 10 affective test did it in only two of the seven; and the grade 8 affective test and the grade 10 cognitive test discriminated against program students in two of the seven. (See Table 20 .)

Validity of the Tests as Measures of Career Education

The most important single question about the tests was whether they measured the objectives of the Ohio Career Development Program. Particularly because the 1973 curriculum-based tests designed to measure intended learning did not discriminate in favor of program students, while the 1975 field-based tests designed to measure actual learning did discriminate in their favor, the following question arose: Were the 1975 field-based tests a valid measure of important career

information and attitudes? That is, did the actual learning demonstrated by the program students on the field-based tests represent valuable learning about careers?

The persons best able to make that judgment by identifying the objectives of the Ohio Career Development Program were the state and local officials responsible for designing and operating it. We asked four of the local Program Directors* to rate each item on the tests as being 1) important, 2) neutral, or 3) unimportant.

Their opinions were sought before they were given the test results.

Their ratings demonstrated that they regarded most of the items on all the tests as important for students to learn. An average of 87 percent of the items on the tests were rated "important" while only 4 percent were rated "unimportant." (See Table 26.)

This seems particularly significant in light of the fact that the sole criterion for selecting items for the final test was their discriminating power. Their validity as measures of important information and attitudes about careers was not considered. Despite this, a representative group of local Program Directors judged a heavy majority of the items as being important for students to learn.

Reliability of the Tests

Using all data from the administration of the final tests, PSE computed split-half reliability coefficients for each test using the

*Representing one item-writing site, one pilot-testing site, and two final testing sites (two of them urban, one of them rural, and one of them suburban).

Kuder-Richardson 20 formula. While the reliability of the tests differed appreciably, with the longer 40-item cognitive tests proving more reliable than the shorter 20-item affective tests, all reached an acceptable level of reliability as shown by the table below.

Kuder-Richardson Measures of Reliability*

<u>Test</u>	<u>Grade</u>	<u>KR-20 Reliability Coefficient</u>
Cognitive	3	.86
Cognitive	6	.88
Cognitive	8	.88
Cognitive	10	.82
Affective	3	.52
Affective	6	.62
Affective	8	.68
Affective	10	.63

A further analysis by PSE using the Spearman-Brown formula indicated that to raise the reliability of the affective tests from an average of about .60 to an average of .86 (the level reached by the cognitive tests) would require affective tests of 80 items--twice the length of the cognitive tests. This indicates that the affective items are not on the average as reliable as the cognitive items. Inasmuch as feelings are more subject to variation from time to time than knowledge, it is, of course, to be expected that an affective test would have to be longer than a cognitive test to be equally reliable.

*See Table 27 for details.

Difficulty of the Tests

Both the local Program Directors and the Division staff expressed concern during the test development activities that the items be made easy enough so that students could demonstrate clearly what they had in fact been learning. PSE did this successfully--partly through item writing and through item selection following the pilot tests. The result is a set of rather easy tests.

The percent of program students answering cognitive test items correctly in grades 3, 6, 8, and 10 ranged from a low of 16% on one item to a high of 94% on another with an average of 63% for all items. (See Table 28.) The usual rule-of-thumb among test developers who want a test to discriminate among students is that an item selected for the test should be answered correctly by 50% of the students.

The four tests were almost equally easy, with the grade 6 and grade 8 tests being just a bit easier than the others (see Table 28).

The Pool of Additional Test Items

PSE developed 683 cognitive test items but used only 160 items to create the final test, leaving a pool of 523 items (unique items plus duplicated items) for possible future use in developing new tests. Similarly, PSE developed 383 affective test items but used only 80 to create the final tests--leaving a pool of 303 (unique items plus duplicated items) for future use.

The items in this pool were rejected for the final tests because other items were better (discriminated more strongly in favor of program

students) when pilot tested in four sites. Of course, some of the items in the pool fell just short of the cut-off point for use in the final tests and are almost as good as those chosen. Thus they merit further attention.

Moreover, items did not discriminate during the final tests in seven sites exactly as they discriminated during the pilot tests in four sites. Some worked well, others not so well, and others poorly. It is likely that some of the pool items--had they been chosen--would have discriminated in the seven cities as well or better than the final test items.

Furthermore, all items in the pool were based on apparent student learning in at least one of the two item-writing sites. The very existence of an item signifies that its content was being taught in one of those two sites. Perhaps technical flaws in the item prevented it from working properly. Eliminating the flaws might make the item serviceable.

Even the items which failed to discriminate in favor of program students or which actually discriminated against them can guide the Division in identifying aspects of the curriculum and teaching that needed to be strengthened or improved.

In short, the pool of additional test items--523 cognitive and 303 affective--represents an important outcome of the study, one that could be quite valuable in improving the final test instruments in the future.

Field-Based Test Development Procedures

PSE undertook the work on the assumption that tests developed in the field rather than in the office, tests to measure what teachers were teaching rather than what program designers planned for them to teach, tests to measure what students were learning rather than what program designers intended them to learn, tests that were classroom-sensitive rather than curriculum-sensitive--such tests could successfully distinguish between program students and non-program students.

That assumption was correct.

City Rankings Confirm Effectiveness of Field-Based Test Development. The best evidence supporting that assumption appears in the ranking of the seven cities. The ranking is in exact correspondence to what PSE knew about what was happening in local classrooms.

The tests worked best (the term "best" means discriminating most strongly in favor of program students) in the cities where the PSE staff 1) spent time on-site studying the local program and 2) designed the tests to match the local programs. The more PSE knew about the city, the more the program students showed their superiority; the less PSE knew about the city, the less the program students showed their superiority.

This is a highly significant finding. It means that PSE has created a method of developing tests that can successfully identify what program students are learning as well as what they are not learning. And those tests can tell in what ways program students are superior to, equal to, or inferior to non-program students.

NEXT STEPS

Based on PSE's experiences during the past four years in evaluating the Ohio Career Development Program, and based on the findings reported here, we offer the following suggestions to the Division as to desirable next steps.

1. Seven Reports to Seven Cities.

The present report is written for the Division and is not designed to be of direct benefit to each of the seven cities administering the final tests. It would be desirable for the Division to authorize a separate report for each city, based on the final test results, pointing out the apparent strengths and weaknesses of each city in cognitive development and affective development in each of the seven developmental areas. Those reports could derive implications from the test results both for local curriculum development and for local staff development. Moreover, the reports could identify those cities apparently doing an especially good job in various developmental areas at various grade levels and could encourage local Program Directors to visit those cities to gain ideas for strengthening their own programs.

2. Applying Statewide Results to Strengthening State Program.

The current report does not include the careful, systematic application of the findings to what the Division might do in the future to improve the program statewide. The Division may wish to authorize a special report relating the findings to such topics as these:

- a. Clarifying the seven developmental areas to sharpen their definition and avoid overlapping content.
- b. Modifying the State Curriculum Guides to produce better student learning in each of the seven developmental areas.
- c. Planning in-service training for local Program Directors and their staffs during 1975-76.

3. Enlarging Pool of Discriminating Items.

While the present tests show favorable results for each of the seven cities, the fact that the cities rank as they do

indicates that the tests should show even more favorable results for each school district if additional items were written on-site in additional districts. Doing this would make the tests more fully representative of the statewide program by accommodating them to significant local variations.

Creating additional discriminating items would have the following advantages:

- a. The affective tests could be strengthened, especially in grade 8 and grade 10 where they do not discriminate strongly in favor of program students.
 - b. Certain developmental areas could be better represented than they are on the present final tests.
 - c. Every district using the improved tests would have a fairer chance to show good results.
 - d. Equivalent forms of the test could be created -- a necessity for performing pre-test/post-test evaluations in the future and a highly desirable way of keeping teachers' attention focused on the broad range of program objectives rather than having them concentrate on teaching the familiar items in the state tests.
4. Developing special tests for each of the Part D sites. The Division wish to consider developing a special test uniquely designed to fit the program in each of the sites so that program students could demonstrate their superior learning. They could do this on tests for their own sites better than they could on statewide tests and better than they could even on a single test developed for all three sites.
5. Developing New Curriculum-based Tests.

It would be desirable to develop a new set of tests representing "ideal" student learning in each of the seven developmental areas in grades 3, 6, 8, and 10. Such tests would constitute deliberately high standards for local programs to achieve. Equivalent forms should be developed, for reasons given in discussing the field-based tests.

Administering both the revised field-based tests and the improved curriculum-based tests in the spring of 1976 would reveal the discrepancy between what students are actually learning and what they should be learning.

Eventually, the field-based tests and the curriculum-based tests would be merged to produce a test with a high standard

of learning, balanced across all seven developmental areas, discriminating in favor of program students, and sensitive to local programs statewide so that each district would have a fair chance to demonstrate its accomplishments.

6. Conducting "Process" Evaluations in Participating Districts.

It would be desirable to repeat the site visits conducted in 1972 to twenty participating school districts (now thirty) to evaluate the processes they are using.

Each district could receive a fall, 1975 report based on early observations recommending how the program should be strengthened during the year. Those recommendations could be confirmed or modified on the basis of results from the revised field-based tests and/or the improved curriculum-based tests administered in spring, 1976 as described earlier.

This would give a powerful combination of information about the relations between local teaching practices and local test results and could lead to a statewide interchange of practices known to produce good test results.

Table 18 (page 1)

Number and Percent of Students
Program and Non-Program

City Code	Grade	Number		Total	Percent	
		Program	Non-program		Program	Non-program
1	3	385	177	562	69	31
	6	391	203	594	66	34
	8	444	208	652	68	32
	10	300	182	482	62	38
2	3	381	143	524	73	27
	6	427	140	567	75	25
	8	393	85	478	82	18
	10	230	117	347	66	34
3	3	368	131	499	74	26
	6	426	129	555	78	22
	8	375	114	489	76	24
	10	341	132	473	72	28
4	3	368	123	455	73	27
	6	385	105	490	79	21
	8	345	129	474	73	27
	10	265	106	371	71	29

Table 18 (page 2)
Number and Percent of Students
Program and Non-Program

City Code	Grade	Number		Total	Percent	
		Program	Non-program		Program	Non-program
5	3	150	128	278	53	47
	6	151	122	273	55	45
	8	194	170	364	53	47
	10	142	163	305	47	53
6	3	359	137	496	72	28
	6	342	138	480	71	29
	8	260	91	351	74	26
	10	242	96	338	72	28
7	3	266	95	361	74	26
	6	302	88	390	77	23
	8	266	86	352	76	24
	10	241	117	358	67	33
State	3	2241	934	3175	71	29
	6	2424	925	3349	72	28
	8	2277	883	3160	72	28
	10	1761	913	2674	66	34

Table 18 (page 3)

Number and Percent of Students

Boys and Girls

City Code	Grade	Number				Percent			
		Program		Non-program		Total		Program	
		Boys	Girls	Boys	Girls	Boys	Girls	Boys	Girls
1	3	177	207	93	83	270	290	16	18
	6	201	186	93	110	294	296	17	16
	8	220	224	206	99	326	323	17	17
	10	153	145	98	82	251	227	16	15
2	3	200	181	78	65	278	246	19	17
	6	194	233	77	62	271	295	17	21
	8	191	202	51	34	242	236	20	21
	10	95	131	72	45	167	176	14	19
3	3	194	174	59	72	253	246	19	17
	6	207	218	62	67	269	285	19	20
	8	169	206	53	61	222	267	17	21
	10	160	180	52	80	212	260	17	19
4	3	149	182	67	55	216	237	16	20
	6	196	189	45	60	241	249	20	19
	8	164	181	59	69	223	250	17	19
	10	118	145	51	53	169	198	16	20

Table 18 (page 4)

Number and Percent of Students

Boys and Girls

City Code	Grade	Number				Percent			
		Program		Non-program		Program		Non-program	
		Boys	Girls	Boys	Girls	Boys	Girls	Boys	Girls
5	3	84	66	62	146	15	12	11	12
	6	85	66	72	157	16	12	13	9
	9	79	115	86	165	11	16	12	11
	10	57	85	93	150	9	14	15	11
6	3	184	175	74	258	19	18	7	6
	6	179	161	71	250	19	17	7	7
	8	111	149	46	157	16	21	7	6
	10	109	128	47	156	14	16	6	6
7	3	114	152	42	156	16	21	6	7
	6	149	150	47	196	19	19	6	5
	8	137	127	47	184	20	18	7	5
	10	130	110	52	182	18	15	7	9
State	3	1102	1137	475	1577	17	18	7	7
	6	1211	1203	467	1678	18	18	7	7
	8	1071	1204	448	1519	17	19	7	7
	10	822	924	465	1287	15	17	9	8

Table 18 (page 5)

Number and Percent of Students

Black, White, Other

City Code	Grade	Number						Percent					
		Program			Non-program			Program			Non-program		
		Black	White	Other	Black	White	Other	Black	White	Other	Black	White	Other
1	3	10	352	23	10	148	19	2	63	4	2	26	3
	6	8	340	43	4	174	25	2	57	7	7	29	4
	8	10	394	40	6	182	20	2	60	6	1	28	3
	10	6	280	14	2	169	11	1	58	3	4	35	2
2	3	316	1	64	17	86	40	60	0	12	3	16	8
	6	390	1	36	27	74	39	69	2	6	5	13	7
	8	359	1	33	7	57	21	75	2	7	1	12	4
	10	217	0	13	6	77	34	63	0	4	2	22	10
3	3	66	268	34	6	99	26	13	54	7	1	20	5
	6	97	279	50	8	108	13	17	50	9	1	19	2
	8	80	263	32	22	82	10	16	54	7	4	17	2
	10	130	173	38	44	77	11	27	37	8	9	16	2
4	3	71	233	28	7	108	8	16	51	6	2	24	2
	6	94	249	42	12	78	15	19	50	9	3	16	3
	8	94	200	51	17	106	6	20	42	11	4	22	1
	10	90	155	20	11	91	4	24	42	5	3	25	1

Table 18 (page 6)

Number and Percent of Students

Black, White, Other

City Code	Grade	Number						Percent					
		Program			Non-program			Program			Non-program		
		Black	White	Other	Black	White	Other	Black	White	Other	Black	White	Other
5	3	0	149	1	1	123	4	0	54	4	4	44	1
	6	1	146	4	1	116	5	4	53	1	4	42	2
	8	1	179	14	0	163	7	3	49	4	0	45	2
	10	0	132	10	2	152	9	0	43	3	6	50	3
6	3	195	126	38	51	72	14	39	25	8	10	15	3
	6	138	172	32	69	49	20	29	36	7	14	10	4
	8	108	138	14	66	18	7	31	39	4	19	5	2
	10	89	136	17	53	25	18	26	40	5	16	7	5
7	3	52	190	24	7	82	6	14	53	7	2	23	2
	6	37	249	16	9	76	3	9	64	4	2	19	18
	8	32	211	23	26	51	9	9	60	7	7	14	3
	10	40	186	15	32	79	6	11	52	4	9	22	2
State	3	710	1319	212	99	718	117	22	42	7	3	23	4
	6	765	1436	223	130	675	120	23	43	7	4	20	4
	8	684	1386	207	144	659	80	22	44	7	5	21	3
	10	527	1062	127	150	670	93	20	40	5	6	25	4

Table 18 (page 7)

Number and Percent of Grade 10 Students

General, Vocational, Academic, Other

City Code	Grade	Number						Percent					
		Program			Non-program			Program			Non-program		
		Gen	Voc	Acad	Other	Gen	Voc	Acad	Other	Gen	Voc	Acad	Other
1	3	-	-	-	-	-	-	-	-	-	-	-	-
	6	-	-	-	-	-	-	-	-	-	-	-	-
	8	-	-	-	-	-	-	-	-	-	-	-	-
	10	113	30	112	45	72	23	65	22	23	6	23	9
										15	5	13	5
2	3	-	-	-	-	-	-	-	-	-	-	-	-
	6	-	-	-	-	-	-	-	-	-	-	-	-
	8	-	-	-	-	-	-	-	-	-	-	-	-
	10	60	69	73	28	22	24	48	23	17	19	21	8
										6	7	14	7
3	3	-	-	-	-	-	-	-	-	-	-	-	-
	6	-	-	-	-	-	-	-	-	-	-	-	-
	8	-	-	-	-	-	-	-	-	-	-	-	-
	10	104	73	125	39	22	78	23	9	22	15	26	8
										5	16	5	2
4	3	-	-	-	-	-	-	-	-	-	-	-	-
	6	-	-	-	-	-	-	-	-	-	-	-	-
	8	-	-	-	-	-	-	-	-	-	-	-	-
	10	91	38	102	34	32	12	46	16	24	10	27	9
										9	3	13	4

Table 18 (page 8)

Number and Percent of Grade 10 Students

General, Vocational, Academic, Other

City Code	Grade	Number						Percent					
		Program			Non-program			Program			Non-program		
		Gen	Voc	Acad	Other	Gen	Voc	Acad	Other	Gen	Voc	Acad	Other
5	3	-	-	-	-	-	-	-	-	-	-	-	-
	6	-	-	-	-	-	-	-	-	-	-	-	-
	8	-	-	-	-	-	-	-	-	-	-	-	-
	10	48	28	59	7	92	16	41	14	16	9	19	2
6	3	-	-	-	-	-	-	-	-	-	-	-	-
	6	-	-	-	-	-	-	-	-	-	-	-	-
	8	-	-	-	-	-	-	-	-	-	-	-	-
	10	112	28	70	52	38	4	44	10	33	8	21	9
7	3	-	-	-	-	-	-	-	-	-	-	-	-
	6	-	-	-	-	-	-	-	-	-	-	-	-
	8	-	-	-	-	-	-	-	-	-	-	-	-
	10	44	18	134	45	48	21	23	25	12	5	37	13
State	3	-	-	-	-	-	-	-	-	-	-	-	-
	6	-	-	-	-	-	-	-	-	-	-	-	-
	8	-	-	-	-	-	-	-	-	-	-	-	-
	10	572	284	675	230	326	178	290	119	20	11	26	9
	3	-	-	-	-	-	-	-	-	-	-	-	-
	6	-	-	-	-	-	-	-	-	-	-	-	-
	8	-	-	-	-	-	-	-	-	-	-	-	-
	10	-	-	-	-	-	-	-	-	-	-	-	-

Table 19

STATE RESULTS

ANALYSIS 1 -- PROGRAM AND NON-PROGRAM RESULTS BY STATE, CITY AND GRADE FOR COGNITIVE AND AFFECTIVE TESTS

(THE SIGNIFICANCE OF THE DIFFERENCE IS SHOWN BY "t"
ANY t LARGER THAN 1.96 IS SIGNIFICANT BEYOND THE .05 LEVEL)

STATE TOTALS		COGNITIVE RESULTS				AFFECTIVE RESULTS			
		Percent Of Items Answered Correctly		Diff	"t"	Percent Of Items Answered Correctly		Diff	"t"
		Program	Non-Program			Program	Non-Program		
GRADE 3	3	59.2	54.1	5.1	16.73	65.7	63.4	2.3	5.56
GRADE 6	6	63.2	59.3	3.8	12.81	77.5	75.9	1.6	4.40
GRADE 8	8	63.5	60.6	2.9	9.60	70.3	69.5	0.8	2.03
GRADE 10	10	61.2	59.4	1.8	5.56	67.6	67.1	0.4	1.04

Table 20 (page 1)

CITY RESULTS

ANALYSIS I -- PROGRAM AND NON-PROGRAM RESULTS BY STATE, CITY AND GRADE FOR COGNITIVE AND AFFECTIVE TESTS

(THE SIGNIFICANCE OF THE DIFFERENCE IS SHOWN BY "t" ANY t LARGER THAN 1.96 IS SIGNIFICANT BEYOND THE .05 LEVEL)

[illegible]

CITY RESULTS -- (CONTINUED)

ANALYSIS I -- PROGRAM AND NON-PROGRAM RESULTS BY STATE, CITY AND GRADE FOR COGNITIVE AND AFFECTIVE TESTS

T-11

Table 21 (page 1)

DEMOGRAPHIC ANALYSIS: MALE

ANALYSIS II-- PROGRAM AND NON-PROGRAM RESULTS BY STATE, CITY AND GRADE FOR COGNITIVE AND AFFECTIVE TESTS

(THE SIGNIFICANCE OF THE DIFFERENCE IS SHOWN BY "t"
ANY t LARGER THAN 1.96 IS SIGNIFICANT BEYOND THE .05 LEVEL)

STATE TOTALS		COGNITIVE RESULTS				AFFECTIVE RESULTS			
		Percent Of Items		Diff	"t"	Percent Of Items		Diff	"t"
		Program	Non-Program			Program	Non-Program		
GRADE 3	3	59.2	53.8	5.4	12.43	65.0	62.4	2.6	4.37
GRADE 6	6	62.4	59.0	3.4	8.14	77.1	74.7	2.3	4.46
GRADE 8	8	61.7	59.5	2.2	5.15	68.4	67.1	1.2	2.12
GRADE 10	10	61.3	58.7	2.6	5.09	66.5	65.8	0.6	1.04

Table 21 (page 2)

DEMOGRAPHIC ANALYSIS: FEMALE

ANALYSIS II-- PROGRAM AND NON-PROGRAM RESULTS BY STATE, CITY AND GRADE FOR COGNITIVE AND AFFECTIVE TESTS

(THE SIGNIFICANCE OF THE DIFFERENCE IS SHOWN BY "t"
ANY t LARGER THAN 1.96 IS SIGNIFICANT BEYOND THE .05 LEVEL)

STATE TOTALS				STATE TOTALS			
COGNITIVE RESULTS				AFFECTIVE RESULTS			
Percent Of Items				Percent Of Items			
Answered Correctly				Answered Correctly			
Program Non-Program				Program Non-Program			
GRADE	3	59.2	Diff	GRADE	3	56.4	Diff
GRADE	6	63.9	4.1	GRADE	6	78.0	1.9
GRADE	8	65.1	3.4	GRADE	8	72.0	0.8
GRADE	10	61.4	1.2	GRADE	10	68.7	0.1
						68.4	0.3
							"t"
							3.27
							1.64
							0.18
							0.50

Table 22 (page 1)

DEMOGRAPHIC ANALYSIS: BLACK

ANALYSIS II-- PROGRAM AND NON-PROGRAM RESULTS BY STATE, CITY AND GRADE FOR COGNITIVE AND AFFECTIVE TESTS

(THE SIGNIFICANCE OF THE DIFFERENCE IS SHOWN BY "t"
ANY t LARGER THAN 1.96 IS SIGNIFICANT BEYOND THE .05 LEVEL)

STATE TOTALS				STATE TOTALS			
COGNITIVE RESULTS				AFFECTIVE RESULTS			
Percent Of Items Answered Correctly				Percent Of Items Answered Correctly			
Program	Non-Program	Diff	"t"	Program	Non-Program	Diff	"t"
GRADE 3	51.1	45.1	6.0	61.7	61.6	0.2	0.13
GRADE 6	54.7	49.5	5.2	75.1	72.0	3.1	3.30
GRADE 8	51.4	47.0	4.3	65.7	62.2	3.6	3.63
GRADE 10	53.0	49.1	3.9	63.9	60.9	3.1	3.07

Table 22 (page 2)

DEMOGRAPHIC ANALYSIS: WHITE

ANALYSIS II-- PROGRAM AND NON-PROGRAM RESULTS BY STATE, CITY AND GRADE FOR COGNITIVE AND AFFECTIVE TESTS

(THE SIGNIFICANCE OF THE DIFFERENCE IS SHOWN BY "t"
ANY t LARGER THAN 1.96 IS SIGNIFICANT BEYOND THE .05 LEVEL)

STATE TOTALS		COGNITIVE RESULTS				AFFECTIVE RESULTS			
		Percent Of Items		Diff	"t"	Percent Of Items		Diff	"t"
		Answered	Correctly			Answered	Correctly		
		Program	Non-Program			Program	Non-Program		
GRADE 3	3	64.1	56.2	7.8	21.73	68.1	64.1	3.9	8.00
GRADE 6	6	68.2	62.2	6.0	17.00	79.1	76.8	2.3	5.33
GRADE 8	8	69.7	64.4	5.3	14.96	72.8	71.6	1.2	2.45
GRADE 10	10	66.0	62.5	3.5	9.55	69.5	69.1	0.4	0.82

Table 22 (page 3)

DEMOGRAPHIC ANALYSIS: OTHER

ANALYSIS II-- PROGRAM AND NON-PROGRAM RESULTS BY STATE, CITY AND GRADE FOR COGNITIVE AND AFFECTIVE TESTS

(THE SIGNIFICANCE OF THE DIFFERENCE IS SHOWN BY "t"
ANY t LARGER THAN 1.96 IS SIGNIFICANT BEYOND THE .05 LEVEL)

STATE TOTALS		COGNITIVE RESULTS				AFFECTIVE RESULTS			
		Percent Of Items Answered Correctly		Diff	"t"	Percent Of Items Answered Correctly		Diff	"t"
		Program	Non-Program			Program	Non-Program		
GRADE 3	3	55.8	48.2	7.6	8.34	64.1	60.1	4.0	3.18
GRADE 6	6	59.8	54.0	5.8	6.50	75.5	75.1	0.4	0.38
GRADE 8	8	62.2	53.4	8.8	8.52	69.0	65.3	3.7	2.67
GRADE 10	10	57.7	54.4	3.2	3.03	67.7	63.0	4.7	3.21
STATE TOTALS									

Table 23 (page 1)
TENTH GRADE CURRICULAR ANALYSIS: GENERAL

ANALYSIS 11-- PROGRAM AND NON-PROGRAM RESULTS BY STATE AND CITY FOR COGNITIVE AND AFFECTIVE TESTS

(THE SIGNIFICANCE OF THE DIFFERENCE IS SHOWN BY "t"
ANY t LARGER THAN 1.96 IS SIGNIFICANT BEYOND THE .05 LEVEL)

COGNITIVE RESULTS					AFFECTIVE RESULTS				
	Percent Of Items Answered Correctly		Diff	"t"		Percent Of Items Answered Correctly		Diff	"t"
	Program	Non-Program				Program	Non-Program		
CITY 1	62.4	56.4	6.0	5.12	CITY 1	66.4	66.7	-0.4	-0.23
CITY 2	51.5	49.0	2.5	1.28	CITY 2	63.0	60.0	3.0	1.10
CITY 3	57.7	54.3	3.4	1.82	CITY 3	67.2	67.0	0.2	0.07
CITY 4	54.3	56.3	-2.0	-1.25	CITY 4	64.5	65.3	-0.9	-0.39
CITY 5	58.8	60.6	-1.8	-1.30	CITY 5	70.3	66.8	3.5	1.88
CITY 6	52.1	41.6	10.5	7.17	CITY 6	62.9	51.8	11.0	5.30
CITY 7	55.4	54.7	0.7	0.40	CITY 7	63.2	66.8	-3.4	-1.52
STATE	56.3	55.0	1.3	2.36	STATE	65.3	64.4	0.8	1.12

Table 23 (page 2)

TENTH GRADE CURRICULAR ANALYSIS: ACADEMIC

ANALYSIS II-- PROGRAM AND NON-PROGRAM RESULTS BY STATE AND CITY FOR COGNITIVE AND AFFECTIVE TESTS

(THE SIGNIFICANCE OF THE DIFFERENCE IS SHOWN BY "t"
ANY t LARGER THAN 1.96 IS SIGNIFICANT BEYOND THE .05 LEVEL)

COGNITIVE RESULTS					AFFECTIVE RESULTS				
	Percent Of Items Answered Correctly		Diff	"t"		Percent Of Items Answered Correctly		Diff	"t"
	Program	Non-Program				Program	Non-Program		
CITY 1	72.2	72.7	-0.5	-0.46	CITY 1	70.2	74.2	-4.0	-2.56
CITY 2	59.7	69.3	-9.7	-6.95	CITY 2	72.4	71.3	1.1	0.61
CITY 3	69.5	69.9	-0.4	-0.21	CITY 3	72.3	75.2	-2.9	-1.33
CITY 4	66.9	74.2	-7.3	-5.78	CITY 4	72.4	75.2	-2.8	-1.62
CITY 5	72.6	72.3	0.4	0.26	CITY 5	76.8	74.4	2.4	1.22
CITY 6	68.6	63.6	5.0	3.47	CITY 6	71.1	68.3	2.8	1.44
CITY 7	72.5	65.3	7.1	4.24	CITY 7	68.9	71.3	-2.4	-1.04
STATE	69.3	70.1	-0.9	-1.69	STATE	71.6	72.8	-1.3	-1.82

Table 23 (page 3)

TENTH GRADE CURRICULAR ANALYSIS: VOCATIONAL

ANALYSIS II-- PROGRAM AND NON-PROGRAM RESULTS BY STATE AND CITY FOR COGNITIVE AND AFFECTIVE TESTS

(THE SIGNIFICANCE OF THE DIFFERENCE IS SHOWN BY "t"
ANY t LARGER THAN 1.96 IS SIGNIFICANT BEYOND THE .05 LEVEL)

COGNITIVE RESULTS

	Percent Of Items Answered Correctly		Diff	"t"	Percent Of Items Answered Correctly		Diff	"t"
	Program	Non-Program			Program	Non-Program		
CITY 1	67.2	59.0	8.1	3.85	70.7	70.2	0.4	0.16
CITY 2	51.8	54.1	-2.2	-1.19	65.1	61.5	3.7	1.44
CITY 3	60.2	55.7	4.6	3.60	67.3	64.9	2.5	1.43
CITY 4	55.4	60.2	-4.8	-1.87	66.7	68.8	-2.0	-0.59
CITY 5	56.8	57.0	-0.2	-0.10	64.6	67.8	-3.2	-0.96
CITY 6	57.1	39.4	17.8	4.30	67.1	51.3	15.9	2.83
CITY 7	57.6	57.6	0.0	0.01	66.9	69.5	-2.6	-0.77
STATE	57.5	56.2	1.3	1.73	66.8	65.9	0.9	0.88

Table 23 (page 4)

TENTH GRADE CURRICULAR ANALYSIS: OTHER

ANALYSIS II-- PROGRAM AND NON-PROGRAM RESULTS BY STATE AND CITY FOR COGNITIVE AND AFFECTIVE TESTS

(THE SIGNIFICANCE OF THE DIFFERENCE IS SHOWN BY "t"
ANY t LARGER THAN 1.96 IS SIGNIFICANT BEYOND THE .05 LEVEL)

COGNITIVE RESULTS					AFFECTIVE RESULTS						
		Percent Of Items Answered Correctly		Diff	"t"			Percent Of Items Answered Correctly		Diff	"t"
		Program	Non-Program			Program	Non-Program				
CITY 1		61.1	46.7	14.4	7.07	CITY 1	67.3	65.2	2.1	0.76	
CITY 2		46.4	46.0	0.5	0.20	CITY 2	56.6	53.7	2.9	0.93	
CITY 3		55.2	45.6	9.6	3.31	CITY 3	63.2	63.3	-0.1	-0.03	
CITY 4		51.9	58.0	-6.1	-2.55	CITY 4	64.3	72.2	-7.9	-2.55	
CITY 5		56.1	55.2	0.9	0.25	CITY 5	72.9	62.7	10.0	2.11	
CITY 6		50.7	52.8	-2.0	-0.72	CITY 6	57.3	57.5	-0.2	-0.04	
CITY 7		56.3	51.7	4.6	2.33	CITY 7	61.4	62.8	-1.4	-0.50	
STATE		54.4	50.5	3.9	4.34	STATE	62.5	62.4	0.1	0.12	

Table 24 (page 1)

DEMOGRAPHIC ANALYSIS: MALE

ANALYSIS II.-- PROGRAM AND NON-PROGRAM RESULTS BY STATE, CITY AND GRADE FOR COGNITIVE AND AFFECTIVE TESTS

(THE SIGNIFICANCE OF THE DIFFERENCE IS SHOWN BY "t"
ANY t LARGER THAN 1.96 IS SIGNIFICANT BEYOND THE .05 LEVEL)

CITY 1				CITY 2				CITY 3				CITY 4			
COGNITIVE RESULTS				COGNITIVE RESULTS				COGNITIVE RESULTS				COGNITIVE RESULTS			
Percent Of Items Answered Correctly				Percent Of Items Answered Correctly				Percent Of Items Answered Correctly				Percent Of Items Answered Correctly			
Program Non-Prog.				Program Non-Prog.				Program Non-Prog.				Program Non-Prog.			
GRADE 3	66.4	59.5	6.9	GRADE 3	68.2	69.0	-0.8	GRADE 3	60.4	58.5	2.0	GRADE 3	63.7	59.7	4.0
GRADE 6	66.0	65.4	0.5	GRADE 6	77.1	74.3	2.8	GRADE 6	74.7	73.2	1.4	GRADE 6	79.7	74.0	5.7
GRADE 8	64.5	64.0	0.5	GRADE 8	65.3	69.6	-4.3	GRADE 8	62.6	66.6	-4.0	GRADE 8	72.0	64.7	7.2
GRADE 10	64.8	58.5	6.3	GRADE 10	67.5	66.9	0.6	GRADE 10	63.6	63.8	-0.2	GRADE 10	67.8	65.5	2.3
									</						

DEMOGRAPHIC ANALYSIS: MALE (CONTINUED)

ANALYSIS II-- PROGRAM AND NON-PROGRAM RESULTS BY STATE, CITY AND GRADE FOR COGNITIVE AND AFFECTIVE TESTS

T-22

DEMOGRAPHIC ANALYSIS: FEMALE

ANALYSIS II-- PROGRAM AND NON-PROGRAM RESULTS BY STATE, CITY AND GRADE FOR COGNITIVE AND AFFECTIVE TESTS

(THE SIGNIFICANCE OF THE DIFFERENCE IS SHOWN BY "t"
ANY t LARGER THAN 1.96 IS SIGNIFICANT BEYOND THE .05 LEVEL)

CITY 1				CITY 2				CITY 3				CITY 4			
COGNITIVE RESULTS				COGNITIVE RESULTS				COGNITIVE RESULTS				COGNITIVE RESULTS			
Percent Of Items Answered Correctly				Percent Of Items Answered Correctly				Percent Of Items Answered Correctly				Percent Of Items Answered Correctly			
Program	Non-Program	Diff	"t"	Program	Non-Program	Diff	"t"	Program	Non-Program	Diff	"t"	Program	Non-Program	Diff	"t"
GRADE 3	59.4	7.7	7.75	GRADE 3	59.6	69.4	72.7	GRADE 3	61.8	60.0	1.8	GRADE 3	67.6	65.3	2.3
GRADE 6	65.4	5.7	6.47	GRADE 6	80.2	77.3	2.8	GRADE 6	75.5	72.3	3.2	GRADE 6	77.9	78.7	-0.8
GRADE 8	66.5	5.6	6.30	GRADE 8	74.1	74.0	0.1	GRADE 8	67.4	66.2	1.2	GRADE 8	73.9	78.7	-3.63
GRADE 10	68.1	3.5	3.37	GRADE 10	69.4	72.7	-2.39	GRADE 10	67.6	63.4	4.2	GRADE 10	68.7	72.1	-3.4
AFFECTIVE RESULTS				AFFECTIVE RESULTS				AFFECTIVE RESULTS				AFFECTIVE RESULTS			
Percent Of Items Answered Correctly				Percent Of Items Answered Correctly				Percent Of Items Answered Correctly				Percent Of Items Answered Correctly			
Program	Non-Program	Diff	"t"	Program	Non-Program	Diff	"t"	Program	Non-Program	Diff	"t"	Program	Non-Program	Diff	"t"
GRADE 3	69.5	0.2	0.13	GRADE 3	69.6	69.4	-2.39	GRADE 3	65.7	63.9	1.8	GRADE 3	67.6	65.3	2.3
GRADE 6	77.3	2.8	2.57	GRADE 6	80.2	77.3	2.8	GRADE 6	77.7	74.1	3.6	GRADE 6	77.9	78.7	-0.8
GRADE 8	74.0	0.1	0.10	GRADE 8	74.1	74.0	0.1	GRADE 8	71.8	64.8	7.0	GRADE 8	73.9	78.7	-3.63
GRADE 10	72.7	-3.3	-2.39	GRADE 10	69.4	72.7	-2.39	GRADE 10	69.4	67.9	1.5	GRADE 10	68.7	72.1	-3.4

Table 24 (page 4)

DEMOGRAPHIC ANALYSIS: FEMALE -- (CONTINUED)

ANALYSIS 11-- PROGRAM AND NON-PROGRAM RESULTS BY STATE, CITY AND GRADE FOR COGNITIVE AND AFFECTIVE TESTS

CITY 5		CITY 5				CITY 6				CITY 7			
GRADE		COGNITIVE RESULTS		Diff	"t"	COGNITIVE RESULTS		Diff	"t"	COGNITIVE RESULTS		Diff	"t"
		Percent Of Items Answered Correctly	Non-Program			Percent Of Items Answered Correctly	Non-Program			Percent Of Items Answered Correctly	Non-Program		
GRADE 3		65.8	60.9	4.8	3.63	53.0	50.9	2.1	1.82	53.8	52.5	1.2	0.96
GRADE 6		70.0	63.3	6.6	4.73	56.4	58.0	-1.5	-1.36	64.9	63.4	1.5	1.14
GRADE 8		70.5	67.3	3.3	3.11	57.7	48.3	9.4	6.98	70.2	59.3	10.9	7.72
GRADE 10		66.6	64.7	1.8	1.52	55.8	50.1	2.6	1.98	66.7	55.9	10.8	8.96
CITY 5		AFFECTIVE RESULTS		Diff	"t"	AFFECTIVE RESULTS		Diff	"t"	AFFECTIVE RESULTS		Diff	"t"
		Percent Of Items Answered Correctly	Non-Program			Percent Of Items Answered Correctly	Non-Program			Percent Of Items Answered Correctly	Non-Program		
GRADE 3		71.1	66.3	4.8	2.65	65.8	61.5	4.2	2.65	65.5	63.1	2.3	1.37
GRADE 6		83.7	83.8	-0.1	-0.06	76.8	78.4	-1.7	-1.23	78.3	76.3	2.0	1.20
GRADE 8		76.9	79.0	-2.1	-1.58	67.5	61.7	5.8	3.17	74.6	67.8	6.8	3.58
GRADE 10		74.5	70.1	4.4	2.71	64.7	60.4	4.3	2.34	68.1	68.1	0.0	0.01
CITY 6		COGNITIVE RESULTS		Diff	"t"	AFFECTIVE RESULTS		Diff	"t"	AFFECTIVE RESULTS		Diff	"t"
		Percent Of Items Answered Correctly	Non-Program			Percent Of Items Answered Correctly	Non-Program			Percent Of Items Answered Correctly	Non-Program		
GRADE 3		53.0	50.9	2.1	1.82	65.8	61.5	4.2	2.65	65.5	63.1	2.3	1.37
GRADE 6		56.4	58.0	-1.5	-1.36	76.8	78.4	-1.7	-1.23	78.3	76.3	2.0	1.20
GRADE 8		57.7	48.3	9.4	6.98	67.5	61.7	5.8	3.17	74.6	67.8	6.8	3.58
GRADE 10		55.8	50.1	2.6	1.98	64.7	60.4	4.3	2.34	68.1	68.1	0.0	0.01
CITY 7		COGNITIVE RESULTS		Diff	"t"	AFFECTIVE RESULTS		Diff	"t"	AFFECTIVE RESULTS		Diff	"t"
		Percent Of Items Answered Correctly	Non-Program			Percent Of Items Answered Correctly	Non-Program			Percent Of Items Answered Correctly	Non-Program		
GRADE 3		53.8	52.5	1.2	0.96	65.5	63.1	2.3	1.37	65.5	63.1	2.3	1.37
GRADE 6		64.9	63.4	1.5	1.14	78.3	76.3	2.0	1.20	78.3	76.3	2.0	1.20
GRADE 8		70.2	59.3	10.9	7.72	74.6	67.8	6.8	3.58	74.6	67.8	6.8	3.58
GRADE 10		66.7	55.9	10.8	8.96	68.1	68.1	0.0	0.01	68.1	68.1	0.0	0.01

DEMOGRAPHIC ANALYSIS: WHITE

ANALYSIS II-- PROGRAM AND NON-PROGRAM RESULTS BY STATE, CITY AND GRADE FOR COGNITIVE AND AFFECTIVE TESTS

(THE SIGNIFICANCE OF THE DIFFERENCE IS SHOWN BY "t"
ANY t LARGER THAN 1.96 IS SIGNIFICANT BEYOND THE .05 LEVEL)

CITY 1				CITY 2				CITY 3				CITY 4			
COGNITIVE RESULTS				COGNITIVE RESULTS				COGNITIVE RESULTS				COGNITIVE RESULTS			
Percent Of Items Answered Correctly				Percent Of Items Answered Correctly				Percent Of Items Answered Correctly				Percent Of Items Answered Correctly			
Program	Non-Program	Diff	"t"	Program	Non-Program	Diff	"t"	Program	Non-Program	Diff	"t"	Program	Non-Program	Diff	"t"
GRADE 3	67.0	7.0	9.33	GRADE 3	69.0			GRADE 3	65.6			GRADE 3	67.8		
GRADE 6	68.4	2.1	3.02	GRADE 6	78.4			GRADE 6	80.2			GRADE 6	78.5		
GRADE 8	69.2	3.3	4.95	GRADE 8	70.3			GRADE 8	73.2			GRADE 8	74.9		
GRADE 10	66.8	5.2	6.99	GRADE 10	68.4			GRADE 10	72.0			GRADE 10	69.8		
AFFECTIVE RESULTS				AFFECTIVE RESULTS				AFFECTIVE RESULTS				AFFECTIVE RESULTS			
Percent Of Items Answered Correctly				Percent Of Items Answered Correctly				Percent Of Items Answered Correctly				Percent Of Items Answered Correctly			
Program	Non-Program	Diff	"t"	Program	Non-Program	Diff	"t"	Program	Non-Program	Diff	"t"	Program	Non-Program	Diff	"t"
GRADE 3	69.3	-0.2	-0.22	GRADE 3	65.0			GRADE 3	65.6			GRADE 3	65.3		
GRADE 6	76.1	2.3	2.58	GRADE 6	80.0			GRADE 6	74.3			GRADE 6	77.9		
GRADE 8	72.3	-2.1	-2.29	GRADE 8	50.0			GRADE 8	64.9			GRADE 8	77.9		
GRADE 10	69.8	-1.3	-1.33	GRADE 10	0.0			GRADE 10	66.9			GRADE 10	71.4		
CITY 1				CITY 2				CITY 3				CITY 4			
COGNITIVE RESULTS				COGNITIVE RESULTS				COGNITIVE RESULTS				COGNITIVE RESULTS			
Percent Of Items Answered Correctly				Percent Of Items Answered Correctly				Percent Of Items Answered Correctly				Percent Of Items Answered Correctly			
Program	Non-Program	Diff	"t"	Program	Non-Program	Diff	"t"	Program	Non-Program	Diff	"t"	Program	Non-Program	Diff	"t"
GRADE 3	67.5	17.6	2.37	GRADE 3	65.0			GRADE 3	65.6			GRADE 3	65.3		
GRADE 6	37.5	-16.0	-2.08	GRADE 6	80.0			GRADE 6	74.3			GRADE 6	77.9		
GRADE 8	50.0	-2.1	-0.26	GRADE 8	50.0			GRADE 8	64.9			GRADE 8	77.9		
GRADE 10	0.0	-62.3	0.00	GRADE 10	0.0			GRADE 10	66.9			GRADE 10	71.4		
CITY 1				CITY 2				CITY 3				CITY 4			
AFFECTIVE RESULTS				AFFECTIVE RESULTS				AFFECTIVE RESULTS				AFFECTIVE RESULTS			
Percent Of Items Answered Correctly				Percent Of Items Answered Correctly				Percent Of Items Answered Correctly				Percent Of Items Answered Correctly			
Program	Non-Program	Diff	"t"	Program	Non-Program	Diff	"t"	Program	Non-Program	Diff	"t"	Program	Non-Program	Diff	"t"
GRADE 3	49.9			GRADE 3	65.0			GRADE 3	65.6			GRADE 3	65.3		
GRADE 6	53.5			GRADE 6	80.0			GRADE 6	74.3			GRADE 6	77.9		
GRADE 8	52.1			GRADE 8	50.0			GRADE 8	64.9			GRADE 8	77.9		
GRADE 10	62.3			GRADE 10	0.0			GRADE 10	66.9			GRADE 10	71.4		
CITY 1				CITY 2				CITY 3				CITY 4			
COGNITIVE RESULTS				COGNITIVE RESULTS				COGNITIVE RESULTS				COGNITIVE RESULTS			
Percent Of Items Answered Correctly				Percent Of Items Answered Correctly				Percent Of Items Answered Correctly				Percent Of Items Answered Correctly			
Program	Non-Program	Diff	"t"	Program	Non-Program	Diff	"t"	Program	Non-Program	Diff	"t"	Program	Non-Program	Diff	"t"
GRADE 3	64.9	11.9	13.00	GRADE 3	65.6			GRADE 3	65.6			GRADE 3	65.3		
GRADE 6	70.1	13.9	16.00	GRADE 6	80.2			GRADE 6	74.3			GRADE 6	77.9		
GRADE 8	59.3	14.3	14.94	GRADE 8	73.2			GRADE 8	64.9			GRADE 8	77.9		
GRADE 10	69.6	9.8	9.41	GRADE 10	72.0			GRADE 10	66.9			GRADE 10	71.4		
CITY 1				CITY 2				CITY 3				CITY 4			
AFFECTIVE RESULTS				AFFECTIVE RESULTS				AFFECTIVE RESULTS				AFFECTIVE RESULTS			
Percent Of Items Answered Correctly				Percent Of Items Answered Correctly				Percent Of Items Answered Correctly				Percent Of Items Answered Correctly			
Program	Non-Program	Diff	"t"	Program	Non-Program	Diff	"t"	Program	Non-Program	Diff	"t"	Program	Non-Program	Diff	"t"
GRADE 3	63.9	4.8	5.38	GRADE 3	67.8			GRADE 3	67.8			GRADE 3	65.3		
GRADE 6	70.9	7.3	7.48	GRADE 6	78.5			GRADE 6	78.5			GRADE 6	77.9		
GRADE 8	66.5	-6.8	-7.85	GRADE 8	74.9			GRADE 8	74.9			GRADE 8	77.9		
GRADE 10	62.4	-3.7	-3.67	GRADE 10	69.8			GRADE 10	69.8			GRADE 10	71.4		
CITY 1				CITY 2				CITY 3				CITY 4			
AFFECTIVE RESULTS				AFFECTIVE RESULTS				AFFECTIVE RESULTS				AFFECTIVE RESULTS			
Percent Of Items Answered Correctly				Percent Of Items Answered Correctly				Percent Of Items Answered Correctly				Percent Of Items Answered Correctly			
Program	Non-Program	Diff	"t"	Program	Non-Program	Diff	"t"	Program	Non-Program	Diff	"t"	Program	Non-Program	Diff	"t"
GRADE 3	69.3	-0.2	-0.22	GRADE 3	65.0			GRADE 3	65.6			GRADE 3	65.3		
GRADE 6	76.1	2.3	2.58	GRADE 6	80.0			GRADE 6	74.3			GRADE 6	77.9		
GRADE 8	72.3	-2.1	-2.29	GRADE 8	50.0			GRADE 8	64.9			GRADE 8	77.9		
GRADE 10	69.8	-1.3	-1.33	GRADE 10	0.0			GRADE 10	66.9			GRADE 10	71.4		
CITY 1				CITY 2				CITY 3				CITY 4			
AFFECTIVE RESULTS				AFFECTIVE RESULTS				AFFECTIVE RESULTS				AFFECTIVE RESULTS			
Percent Of Items Answered Correctly				Percent Of Items Answered Correctly				Percent Of Items Answered Correctly				Percent Of Items Answered Correctly			
Program	Non-Program	Diff	"t"	Program	Non-Program	Diff	"t"	Program	Non-Program	Diff	"t"	Program	Non-Program	Diff	"t"
GRADE 3	69.3	-0.2	-0.22	GRADE 3	65.0			GRADE 3	65.6			GRADE 3	65.3		
GRADE 6	76.1	2.3	2.58	GRADE 6	80.0			GRADE 6	74.3			GRADE 6	77.9		
GRADE 8	72.3	-2.1	-2.29	GRADE 8	50.0			GRADE 8	64.9			GRADE 8	77.9		
GRADE 10	69.8	-1.3	-1.33	GRADE 10	0.0			GRADE 10	66.9			GRADE 10	71.4		
CITY 1				CITY 2				CITY 3				CITY 4			
AFFECTIVE RESULTS				AFFECTIVE RESULTS				AFFECTIVE RESULTS				AFFECTIVE RESULTS			
Percent Of Items Answered Correctly				Percent Of Items Answered Correctly				Percent Of Items Answered Correctly				Percent Of Items Answered Correctly			
Program	Non-Program	Diff	"t"	Program	Non-Program	Diff	"t"	Program	Non-Program	Diff	"t"	Program	Non-Program	Diff	"t"
GRADE 3	69.3	-0.2	-0.22	GRADE 3	65.0			GRADE 3	65.6			GRADE 3	65.3		
GRADE 6	76.1	2.3	2.58	GRADE 6	80.0			GRADE 6	74.3			GRADE 6	77.9		
GRADE 8	72.3	-2.1	-2.29	GRADE 8	50.0			GRADE 8	64.9			GRADE 8	77.9		
GRADE 10	69.8	-1.3	-1.33	GRADE 10	0.0			GRADE 10	66.9			GRADE 10	71.4		
CITY 1				CITY 2				CITY 3				CITY 4			
AFFECTIVE RESULTS				AFFECTIVE RESULTS				AFFECTIVE RESULTS				AFFECTIVE RESULTS			
Percent Of Items Answered Correctly				Percent Of Items Answered Correctly				Percent Of Items Answered Correctly				Percent Of Items Answered Correctly			
Program	Non-Program	Diff	"t"	Program	Non-Program	Diff	"t"	Program	Non-Program	Diff	"t"	Program	Non-Program	Diff	"t"
GRADE 3	69.3	-0.2	-0.22	GRADE 3	65.0			GRADE 3	65.6			GRADE 3	65.3		
GRADE 6	76.1	2.3	2.58	GRADE 6	80.0			GRADE 6	74.3			GRADE 6	77.9		
GRADE 8	72.3	-2.1	-2.29	GRADE 8	50.0			GRADE 8	64.9			GRADE 8	77.9		
GRADE 10	69.8	-1.3	-1.33	GRADE 10	0.0			GRADE 10	66.9			GRADE 10	71.4		
CITY 1				CITY 2				CITY 3				CITY 4			
AFFECTIVE RESULTS				AFFECTIVE RESULTS				AFFECTIVE RESULTS				AFFECTIVE RESULTS			
Percent Of Items Answered Correctly				Percent Of Items Answered Correctly				Percent Of Items Answered Correctly				Percent Of Items Answered Correctly			
Program	Non-Program	Diff	"t"	Program	Non-Program	Diff	"t"	Program	Non-Program	Diff	"t"	Program	Non-Program	Diff	"t"
GRADE 3	69.3	-0.2	-0.22	GRADE 3	65.0			GRADE 3	65.6			GRADE 3	65.3		
GRADE 6	76.1	2.3	2.58	GRADE 6	80.0			GRADE 6	74.3			GRADE 6	77.9		
GRADE 8	72.3	-2.1	-2.29	GRADE 8	50.0			GRADE 8	64.9			GRADE 8	77.9		
GRADE 10	69.8	-1.3	-1.33	GRADE 10	0.0			GRADE 10	66.9			GRADE 10	71.4		
CITY 1				CITY 2				CITY 3				CITY 4			
AFFECTIVE RESULTS				AFFECTIVE RESULTS				AFFECTIVE RESULTS				AFFECTIVE RESULTS			
Percent Of Items Answered Correctly				Percent Of Items Answered Correctly				Percent Of Items Answered Correctly				Percent Of Items Answered Correctly			
Program	Non-Program	Diff	"t"	Program	Non-Program	Diff	"t"	Program	Non-Program	Diff	"t"	Program	Non-Program	Diff	"t"
GRADE 3	69.3	-0.2	-0.22	GRADE 3	65.0			GRADE 3	65.6			GRADE 3	65.3		
GRADE 6	76.1	2.3	2.58	GRADE 6	80.0			GRADE 6	74.3			GRADE 6	77.9		
GRADE 8	72.3	-2.1	-2.29	GRADE 8	50.0			GRADE 8	64.9			GRADE 8	77.9		
GRADE 10	69.8	-1.3	-1.33	GRADE 10	0.0			GRADE 10	66.9			GRADE 10	71.4		
CITY 1				CITY 2				CITY 3				CITY 4			
AFFECTIVE RESULTS				AFFECTIVE RESULTS				AFFECTIVE RESULTS				AFFECTIVE RESULTS			
Percent Of Items Answered Correctly				Percent Of Items Answered Correctly				Percent Of Items Answered Correctly				Percent Of Items Answered Correctly			
Program	Non-Program	Diff	"t"	Program	Non-Program	Diff	"t"	Program	Non-Program	Diff	"t"	Program	Non-Program	Diff	"t"
GRADE 3	69.3	-0.2	-0.22	GRADE 3	65.0			GRADE 3	65.6			GRADE 3	65.3		
GRADE 6	76.1	2.3	2.58	GRADE 6	80.0			GRADE 6	74.3			GRADE 6	77.9		
GRADE 8	72.3	-2.1	-2.29	GRADE 8	50.0			GRADE 8	64.9			GRADE 8	77.9		
GRADE 10	69.8	-1.3	-1.33	GRADE 10	0.0			GRADE 10	66.9			GRADE 10	71.4		
CITY 1				CITY 2				CITY 3				CITY 4			
AFFECTIVE RESULTS				AFFECTIVE RESULTS				AFFECTIVE RESULTS				AFFECTIVE RESULTS			
Percent Of Items Answered Correctly				Percent Of Items Answered Correctly				Percent Of Items Answered Correctly				Percent Of Items Answered Correctly			
Program	Non-Program	Diff	"t"	Program	Non-Program	Diff	"t"	Program	Non-Program	Diff	"t"	Program	Non-Program	Diff	"t"
GRADE 3	69.3	-0.2	-0.22	GRADE 3	65.0			GRADE 3	65.6			GRADE 3	65.3		
GRADE 6	76.1	2.3	2.58	GRADE 6	80.0			GRADE 6	74.3			GRADE 6	77.9		
GRADE 8	72.3	-2.1	-2.29	GRADE 8	50.0			GRADE 8	64.9			GRADE 8	77.9		
GRADE 10	69.8	-1.3	-1.33	GRADE 10	0.0			GRADE 10	66.9			GRADE 10	71.4		
CITY 1				CITY 2				CITY 3				CITY 4			
AFFECTIVE RESULTS				AFFECTIVE RESULTS				AFFECTIVE RESULTS				AFFECTIVE RESULTS			
Percent Of Items Answered Correctly				Percent Of Items Answered Correctly				Percent Of Items Answered Correctly				Percent Of Items Answered Correctly			
Program	Non-Program	Diff	"t"	Program	Non-Program	Diff	"t"	Program	Non-Program	Diff	"t"	Program	Non-Program	Diff	"t"
GRADE 3	69.3	-0.2	-0.22	GRADE 3	65.0			GRADE 3	65.6			GRADE 3	65.3		
GRADE 6	76.1	2.3	2.58	GRADE 6	80.0			GRADE 6	74.3			GRADE 6	77.9		
GRADE 8	72.3	-2.1	-2.29	GRADE 8	50.0			GRADE 8	64.9			GRADE 8	77.9		
GRADE 10	69.8	-1.3	-1.33	GRADE 10	0.0			GRADE 10	66.9			GRADE 10	71.4		
CITY 1				CITY 2				CITY 3				CITY 4			
AFFECTIVE RESULTS				AFFECTIVE RESULTS				AFFECTIVE RESULTS				AFFECTIVE RESULTS			
Percent Of Items Answered Correctly				Percent Of Items Answered Correctly				Percent Of Items Answered Correctly				Percent Of Items Answered Correctly			
Program	Non-Program	Diff	"t"	Program	Non-Program	Diff	"t"	Program	Non-Program	Diff	"t"	Program	Non-Program	Diff	"t"
GRADE 3	69.3	-0.2	-0.22	GRADE 3	65.0			GRADE 3	65.6			GRADE 3	65.3		
GRADE 6	76.1	2.3	2.58	GRADE 6	80.0			GRADE 6	74.3			GRADE 6	77.9		
GRADE 8	72.3	-2.1	-2.29	GRADE 8	50.0			GRADE 8	64.9			GRADE 8	77.9		
GRADE 10	69.8	-1.3	-1.33	GRADE 10	0.0			GRADE 10	66.9			GRADE 10	71.4		
CITY 1				CITY 2				CITY 3				CITY 4			
AFFECTIVE RESULTS				AFFECTIVE RESULTS				AFFECTIVE RESULTS				AFFECTIVE RESULTS			
Percent Of Items Answered Correctly				Percent Of Items Answered Correctly				Percent Of Items Answered Correctly				Percent Of Items Answered Correctly			
Program	Non-Program	Diff	"t"	Program	Non-Program	Diff	"t"	Program	Non-Program	Diff	"t"	Program	Non-Program	Diff	"t"
GRADE 3	69.3	-0.2	-0.22	GRADE 3	65.0			GRADE 3	65.6			GRADE 3	65.3		
GRADE 6	76.1	2.3	2.58	GRADE 6	80.0			GRADE 6	74.3			GRADE 6	77.9		
GRADE 8	72.3	-2.1	-2.29	GRADE 8	50.0			GRADE 8	64.9			GRADE 8	77.9		
GRADE 10	69.8	-1.3	-1.33	GRADE 10	0.0			GRADE 10	66.9			GRADE 10	71.4		
CITY 1				CITY 2				C							

Table 25 (page 2)

DEMOGRAPHIC ANALYSIS: WHITE

-- (CONTINUED)

ANALYSIS II-- PROGRAM AND NON-PROGRAM RESULTS BY STATE, CITY AND GRADE FOR COGNITIVE AND AFFECTIVE TESTS

CITY 5		COGNITIVE RESULTS				AFFECTIVE RESULTS			
		Percent Of Items		Answered Correctly		Percent Of Items		Answered Correctly	
		Program	Non-Prog.	Program	Non-Prog.	Program	Non-Prog.	Program	Non-Prog.
GRADE 3		65.8	57.4	8.4	9.01	70.9	63.7	7.2	5.61
GRADE 6		66.7	62.9	3.8	4.03	82.1	79.3	2.8	2.53
GRADE 8		68.4	65.5	2.9	3.62	76.2	75.4	0.8	0.75
GRADE 10		63.9	63.2	0.7	0.73	71.8	68.9	2.9	2.39
				Diff	"t"			Diff	"t"
				8.4	9.01			7.2	5.61
				3.8	4.03			2.8	2.53
				2.9	3.62			0.8	0.75
				0.7	0.73			2.9	2.39

DEMOGRAPHIC ANALYSIS: BLACK

ANALYSIS II-- PROGRAM AND NON-PROGRAM RESULTS BY STATE, CITY AND GRADE FOR COGNITIVE AND AFFECTIVE TESTS

(THE SIGNIFICANCE OF THE DIFFERENCE IS SHOWN BY "t"
ANY t LARGER THAN 1.96 IS SIGNIFICANT BEYOND THE .05 LEVEL)

CITY 1				CITY 2				CITY 3				CITY 4			
COGNITIVE RESULTS		AFFECTIVE RESULTS		COGNITIVE RESULTS		AFFECTIVE RESULTS		COGNITIVE RESULTS		AFFECTIVE RESULTS		COGNITIVE RESULTS		AFFECTIVE RESULTS	
Grade	Percent Of Items Answered Correctly	Diff	"t"	Grade	Percent Of Items Answered Correctly	Diff	"t"	Grade	Percent Of Items Answered Correctly	Diff	"t"	Grade	Percent Of Items Answered Correctly	Diff	"t"
GRADE 3	53.5	-1.6	-0.50	GRADE 3	52.5	9.1	4.69	GRADE 3	52.6	5.1	1.51	GRADE 3	58.2	11.1	3.56
GRADE 6	61.3	7.5	1.57	GRADE 6	55.6	12.1	7.77	GRADE 6	59.5	10.4	3.58	GRADE 6	58.2	-2.3	-0.95
GRADE 8	56.5	9.8	2.42	GRADE 8	49.5	7.7	2.60	GRADE 8	60.6	15.1	8.00	GRADE 8	52.6	0.5	0.22
GRADE 10	56.3	-3.8	-0.59	GRADE 10	53.9	11.8	3.64	GRADE 10	54.3	1.0	0.69	GRADE 10	53.3	-0.3	-0.13

- - (CONTINUED)

DEMOGRAPHIC ANALYSIS: BLACK

ANALYSIS II-- PROGRAM AND NON-PROGRAM RESULTS BY STATE, CITY AND GRADE FOR COGNITIVE AND AFFECTIVE TESTS

T-28

DEMOGRAPHIC ANALYSIS: OTHER

ANALYSIS II-- PROGRAM AND NON-PROGRAM RESULTS BY STATE, CITY AND GRADE FOR COGNITIVE AND AFFECTIVE TESTS

(THE SIGNIFICANCE OF THE DIFFERENCE IS SHOWN BY "t"
ANY t LARGER THAN 1.96 IS SIGNIFICANT BEYOND THE .05 LEVEL)

CITY	GRADE	COGNITIVE RESULTS Percent Of Items Answered Correctly Program Non-Prog.	Diff	"t"	AFFECTIVE RESULTS Percent Of Items Answered Correctly Program Non-Prog.	Diff	"t"
CITY 1	GRADE 3	63.9	9.0	3.77	70.4	1.5	0.47
	GRADE 6	68.7	8.5	4.45	78.8	3.8	1.61
	GRADE 8	62.3	-0.1	-0.06	66.1	-3.1	-1.10
	GRADE 10	61.6	3.7	1.17	67.1	-0.6	-0.14
CITY 2	GRADE 3	49.9	9.1	5.79	60.8	2.8	1.26
	GRADE 6	58.0	18.1	10.08	73.5	0.9	0.40
	GRADE 8	49.3	10.4	4.78	62.9	-2.1	-0.71
	GRADE 10	48.8	-1.4	-0.56	64.2	6.9	1.95
CITY 3	GRADE 3	61.0	12.7	6.23	65.0	4.8	1.70
	GRADE 6	55.8	9.3	3.83	77.4	3.2	1.05
	GRADE 8	75.2	16.2	5.93	73.8	8.3	2.18
	GRADE 10	56.4	1.4	0.54	67.2	-6.4	-1.87
CITY 4	GRADE 3	59.9	12.7	4.04	66.1	0.4	0.10
	GRADE 6	54.8	-4.8	-2.07	76.0	-0.7	-0.25
	GRADE 8	63.2	-13.8	-4.75	74.2	0.0	0.01
	GRADE 10	59.5	-7.8	-1.88	67.5	-6.3	-1.15

Table 25 (page 6)

.. (CONTINUED)

GRADE FOR COGNITIVE AND AFFECTIVE TESTS

CITY 5		CITY 6		CITY 7	
COGNITIVE RESULTS Percent Of Items Answered Correctly Program Non-Prod.		COGNITIVE RESULTS Percent Of Items Answered Correctly Program Non-Prod.		COGNITIVE RESULTS Percent Of Items Answered Correctly Program Non-Prod.	
GRADE 3	87.5	GRADE 3	49.7	GRADE 3	59.8
GRADE 6	56.9	GRADE 6	50.5	GRADE 6	50.9
GRADE 8	57.5	GRADE 8	66.4	GRADE 8	60.3
GRADE 10	65.8	GRADE 10	57.1	GRADE 10	58.8
Diff	31.3	Diff	-3.7	Diff	-0.6
"t"	4.78	"t"	-1.48	"t"	-0.18
	-1.28		-6.10		1.33
	4.87		4.08		1.10
	17.5		14.6		3.4
	12.7		4.7		-11.6
AFFECTIVE RESULTS Percent Of Items Answered Correctly Program Non-Prod.		AFFECTIVE RESULTS Percent Of Items Answered Correctly Program Non-Prod.		AFFECTIVE RESULTS Percent Of Items Answered Correctly Program Non-Prod.	
GRADE 3	80.0	GRADE 3	62.8	GRADE 3	64.8
GRADE 6	71.3	GRADE 6	71.6	GRADE 6	73.4
GRADE 8	64.3	GRADE 8	69.3	GRADE 8	67.0
GRADE 10	75.0	GRADE 10	65.3	GRADE 10	70.3
Diff	18.8	Diff	10.6	Diff	8.1
"t"	1.79	"t"	3.07	"t"	1.62
	0.33		-3.15		-1.83
	0.57		2.01		2.02
	2.92		1.30		0.07

Table 26

Importance of Test Items
as Rated by Local Project Directors

<u>Grade</u>	<u>Local Program Director</u>	<u>Percent of Test Items Given Rating</u>		
		<u>Important</u>	<u>Neutral</u>	<u>Unimportant</u>
3	City 2	80	18	2
	City 5	88	5	6
	City 6	96	3	0
	City 7	95	5	0
	Average	90	8	2
6	City 2	81	13	5
	City 5	68	28	3
	City 6	95	5	0
	City 7	90	5	5
	Average	84	13	3
8	City 2	90	2	8
	City 5	87	5	8
	City 6	90	8	2
	City 7	75	22	3
	Average	86	9	5
10	City 2	90	10	0
	City 5	81	3	15
	City 6	95	3	2
	City 7	83	13	3
	Average	87	7	5
All Cities Combined Average		87	9	4

Table 27

Kuder-Richardson Measures of Reliability

MEASURE	Cognitive Tests				Affective Tests			
	GRADES							
	3	6	8	10	3	6	8	10
Mean & Score	58	62	63	61	65	77	70	67
Mean Raw Score	23	25	25	24	13	15	14	13
Std. Dev.	7.2	7.7	7.7	6.3	2.8	2.8	3.2	3.0
Variance	51.6	59.5	59.4	40.1	7.7	7.8	10.4	9.3
Alpha	.856	.879	.877	.819	.501	.617	.672	.623

Table 28 (page 1)

Frequency Distribution of Responses

State Totals* for Cognitive Items

Final Tests

Grade 3				Grade 6			
Item No.	Percent **	Item No.	Percent	Item No.	Percent	Item No.	Percent
16	87.8	15	61.6	2	86.0	11	62.5
34	83.9	7	58.6	9	85.4	28	62.2
11	81.0	40	57.4	27	83.5	36	60.4
36	79.0	27	57.0	39	82.7	30	59.1
13	78.4	29	56.4	29	82.5	34	50.9
2	78.4	31	55.5	35	81.9	19	57.8
9	76.3	21	54.6	8	77.7	12	57.7
30	75.4	35	51.1	16	76.2	4	57.5
19	73.8	26	50.7	21	76.1	37	57.6
10	73.6	38	49.7	31	73.8	20	53.6
24	72.9	23	49.5	5	73.7	22	53.0
17	72.4	28	45.4	24	73.6	26	52.4
12	72.3	4	39.0	1	72.8	26	52.4
32	71.9	25	38.2	10	71.5	18	48.1
1	69.4	14	35.1	7	71.3	23	43.2
5	67.2	33	32.9	32	69.9	25	42.8
6	65.2	22	32.2	3	68.4	17	40.8
8	65.1	37	30.6	13	67.1	14	36.3
39	63.4	18	24.3	33	65.3	40	30.0
3	61.8	20	17.9	38	64.5	6	28.8
Mean = 61.7				Mean = 63.5			

* State totals = data for seven cities combined.

** Percent of program students who chose the correct response.

Table 28 (page 2)

Frequency Distribution of Responses

State Totals* for Cognitive Items

Final Tests

Grade 8				Grade 10			
Item No.	Percent **	Item No.	Percent	Item No.	Percent	Item No.	Percent
1	91.7	26	64.1	22	93.9	20	61.6
7	84.7	35	62.3	11	90.3	29	61.4
37	84.6	18	61.8	3	89.3	25	59.5
6	81.9	34	61.4	38	82.9	34	58.1
11	78.0	28	61.0	15	81.9	17	54.6
30	77.7	36	60.1	33	81.8	23	52.9
29	75.7	4	60.1	4	81.4	7	50.3
40	74.4	31	57.5	1	79.4	9	50.0
39	73.3	3	56.4	26	78.8	5	49.3
15	72.3	16	54.0	12	78.4	40	46.2
19	72.2	12	52.9	16	77.6	6	45.8
9	71.7	24	52.7	24	77.1	39	44.5
13	71.6	33	52.6	32	74.6	19	42.4
2	70.1	32	51.9	30	74.0	2	42.0
14	70.0	20	51.4	21	71.2	36	42.3
10	69.9	17	49.1	31	69.2	35	39.3
25	69.2	5	45.6	28	68.2	14	35.5
8	67.4	28	38.9	8	68.0	37	27.8
21	66.4	27	37.4	18	64.1	13	25.0
22	66.3	23	20.9	27	62.4	10	15.8
Mean = 65.2				Mean = 62.0			

* State totals = data for seven cities combined

** Percent of program students who chose the correct response.

Table 28 (page 3)

Frequency Distribution of Responses
State Totals* for Affective Items
Final Tests

Grade 3			Grade 6		Grade 8		Grade 10	
Item	No.	Percent **	Item	No.	Percent	Item	No.	Percent
7		90.6	13		95.8	10		91.6
8		89.1	14		93.3	14		91.2
20		85.6	10		89.9	5		85.9
18		79.7	7		88.9	1		84.8
15		77.3	8		88.3	19		83.1
1		76.9	16		88.2	6		80.2
17		75.5	11		84.7	4		78.6
3		73.5	17		83.2	7		78.0
14		69.4	5		82.5	20		76.5
13		67.2	1		81.5	13		76.4
16		66.6	18		80.7	16		75.8
5		65.8	12		80.2	4		73.6
9		59.8	4		74.4	15		69.8
12		59.0	20		73.1	17		65.5
4		54.1	15		71.1	18		61.6
6		53.5	3		68.2	9		61.1
11		53.4	2		64.8	11		59.7
10		43.2	9		55.4	2		58.8
2		42.6	6		54.7	12		39.6
19		30.9	19		52.0	8		36.1
Mean =	66.9		Mean =	81.6		Mean =	76.1	
						Mean =	69.1	

* State totals = data for seven cities combined

** Percent of program students who chose the correct response.